FUNCTIONAL FAMILY THERAPY: ANALYSIS OF EFFECT AND IMPLEMENTATION

by

CLIO BELLE WEISMAN

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF SOCIAL WORK

School of Social Policy
University of Birmingham
August 2019

ABSTRACT

Purpose: Functional Family Therapy (FFT) is a manualized intervention for youth aged 10-18 with behavioural problems (e.g. delinquency, violence, substance abuse, and truancy) and their families. This thesis examined the evidence surrounding FFT using an overview and an implementation analysis.

Methods: The overview used Cochrane Guidelines as a framework, and included a multi-pronged, highly sensitive search strategy. The narrative implementation analysis was based on the Oxford Implementation Index, and examined the effects of dose, delivery, uptake, context and biases. A reflexive discussion of the research process highlighted the presence of allegiance bias and possibly ethical misconduct by FFT developers.

Results: The overview included 31 reviews and demonstrated effects of the intervention on core outcomes (recidivism and substance abuse) were modest and out-of-home placement was not reported. Secondary outcomes were also modest but generally positive. The implementation analysis included 16 studies involving 5320 participants. Improved training and supervision were associated with better core outcomes; class and ethnicity were important effect moderators whereas there was no apparent dose relationship.

Conclusions: The overall quality of included reviews was low, which makes evidence concerning FFT inconclusive. Implementation elements were shown to affect core outcomes and the likelihood of allegiance bias is very high.

DEDICATION

For Dolly

ACKNOWLEDGEMENTS

Above all, I wish to thank my most excellent supervisor and friend Paul Montgomery, without whom this entire undertaking would have been impossible. That his standards are so high served at each step to make my work better in every regard. He sees the completion of a doctorate as a collaborative effort, and he is absolutely correct. Working with him has been an unmitigated pleasure.

Harry Ferguson, Jane Dennis, Julia Littell, and Gregory Norfleet all contributed to the academic rigour of this project through their feedback, coding efforts, insight and academic knowledge. This thesis was beautifully copy edited for conventions of language, spelling and grammar by Jacoby Patterson

Temple, Jessie, Bunny, Jonny Boy, Avra, Hank…over more than twenty years your friendship has sustained me and perhaps saved me in more ways than I can count. David, you brought beauty into my life over a decade ago, and you continue to show me a world filled with light.

My two moms have maintained an unshakeable conviction that I could in fact do this, and whenever I doubted myself always reminded me to believe them - because we should always listen to our mothers.

The intellectual curiosity instilled in me since childhood I must thank my father for, it has been a tremendous gift.

Contents

TABLE OF FIGURES

1 BACKGROUND

1.1 Outline of Chapter

This chapter begins with a description of behavioural problems in young people age 10 to 18, and the prevalence and impact child behavioural problems have on the individual, families and larger society. The aetiology and risks related to the development of behavioural problems is outlined, and the diagnostic criterion and classification of this group of disorders discussed. Additionally, this section delineates likely reasons a young person will be referred to treatment, and what the most common interventions are in use today for the reduction or prevention of youth behavioural problems. The chapter then goes on to examine in brief some common approaches to the treatment of behavioural problems with a focus on how these issues are addressed in the United States. Next, the history and development of systemic, strategic and structural family therapy models is discussed, as these models form the base for Functional Family Therapy (FFT), which is the model of treatment that is analysed throughout this thesis. FFT is introduced and explained in detail, and finally the intervention's mechanism for change is described.

While behavioural problems in youth are of global concern, this research is predominately located in the United States, due in large part to the fact that FFT was designed and developed in the US, and the vast majority of the trials, studies and secondary research have taken place in an American context.

1.2 Behavioural Problems in Youth

The term 'Behavioural problems', sometimes referred to in the literature as 'serious emotional or behavioural disturbance' is expansive, not a technical diagnosis, and refers to a wide array of psychiatric disorders and psychosocial problems, both internalising and externalising, which includes psychiatric diagnoses, school based behaviours, criminal actions, antisocial

patterns of behaviours and attention deficit problems. These disorders are found in the Diagnostic and Statistical Manual of Mental Disorders: Fifth Edition – Text Revision (DSM-V-TR; American Psychiatric Association, 2013), and are classified as either behavioural or emotional in nature. The ICD-10 Classification of Mental and Behavioural Disorders (ICD-10, WHO) also provides a classification system which is used internationally and was developed by the World Health Organization in conjunction with a number of international agencies including the Alcohol, Drug Abuse and Mental Health Administration (ADAMHA) in the US. This system also separates between organic and behavioural problems and differentiates between those diagnoses that apply to children and youth as opposed to those with a later onset. For diagnostic purposes, the DSM-V is more commonly used, as it provides clearer frameworks and guidance relating to differential diagnosis.

1.2.1 Diagnoses and behaviours

For diagnostic purposes, the DSM-V and ICD-10-CM group behavioural problems into separate categories; Oppositional Defiant Disorder (ODD) and the (usually) later onset and more severe Conduct Disorder (CD), are both grouped under the heading of 'Disruptive, Impulse-Control and Conduct Disorders', a category which also includes Kleptomania, Pyromania and Intermittent Explosive Disorder. While some have attempted to describe ODD as merely a less severe version of CD, this is a fallacy, the two disorders should be treated as separate and discrete. ODD has been shown to be more closely related to inconsistencies and failings in emotional regulation which lead to irritability, aggression and disruptive behaviours, while CD is identified with often limited prosocial emotions, a lack of remorse, lack of empathy and a pervasive pattern of overtly aggressive, deceitful, destructive, violent or cruel actions. Attention-Deficit/Hyperactivity Disorder (ADHD), a diagnosis which often occurs co-morbidly with ODD or CD, is categorized by the DSM-V as a Neurodevelopmental Disorder and is approached and

treated separately and often by different methods. Certain types of sexual acting out, or behavioural problems which do not necessarily warrant the diagnosis of either ODD or CD are categorized under 'Other Conditions that May Be a Focus of Clinical Attention'. This heading is often used to draw the attention of practitioners to the presenting problem, without the necessity of labelling the youth as defiant (DSM-V-TR; American Psychiatric Association, 2013).

For the individual, these disorders are associated with an increase in both internalising and externalising behaviours. Internalizing problems are associated with a lack of control or ability to recognize and cope with emotions, social isolation or withdrawal and lability (McCulloch, Wiggins, Joshi, & Sachdev, 2000). Externalising behaviours are for more likely to bring the attention of authority or parental figures, as they are overt in nature, and may be hostile, aggressive or destructive (Thapar et al., 2015). Specifically, disruptive, impulse control or conduct related disorders directly relate to actions that are conflictive and in violation of other's rights. These disorders are also referred to as disinhibition and often have a high level of comorbidity with substance abuse and later manifestations of more severe personality disorders. Several studies have found that there is comorbidity between internalizing and externalizing behaviour problems (Bornstein, Hahn, & Haynes, 2010; McWey, Cui, & Pazdera, 2010) which may lead to increased chance of antisocial behaviours, self-harm and suicide.

Antisocial behaviours are broadly defined by the DSM V as disruptive actions which may be either covert or overtly hostile and/or aggressive. In order for a behaviour to be considered antisocial, it needs to be more than outside the standards and realm of socially acceptable behaviour, it must contain aggression and hostility (Rutter, Giller, & Hagell 1998; Thapar et al., 2015). Antisocial behaviour must exhibit the intent to do harm, but often antisocial behaviour is not prosecuted or identified as specifically breaking a law. For behaviour to be considered and

categorized as antisocial, it must include elements of either ODD or CD, which contain within them externalising actions against or in the violation of the rights of others (Muncie, 2015; Thapar et al., 2015). These behaviours range in severity and may include defiance, theft, deceitfulness, persistent avoidance of societal norms or rules, recklessness and disregard for the rights of others. Overt behaviours include open aggression, violence, physical or verbal abuse of siblings or peers, destruction of property, sexual misconduct and fire starting. Covert antisocial behaviour may include secretly destroying another's property, lying, or other forms of deception (DSM V; Thapar et al., 2015).

Also falling under the umbrella of antisocial behaviours are more extreme and criminal acts perpetrated by minors, these include the buying or selling of drugs, violent assault, the harming of animals or others, rape, sexual assault, the use or possession of firearms and other criminal behaviours. These types of actions exist further along the same spectrum, and often when these actions come to the attention of law enforcement, it is usually after early warning signs or less destructive and violent behaviours were overlooked by parents, teachers, or counsellors.

When these early warning signs are overlooked or dismissed as normative, a phase, or otherwise unimportant, there is a far greater likelihood that antisocial personality disorder (DSM -V code 301.7; ICD-10-CM code F60.2) may develop, which is a significantly more severe and destructive diagnosis. This personality disorder, which cannot be given until after the age of 18, is defined as "a pervasive pattern of disregard for, and violation of, the rights of others that begins in childhood or early adolescence and continues into adulthood" (p. 659). This is a strong recommendation for early and effective intervention, when a child or youth's externalizing first come to the attention of any potential referral source or other adult. If left untreated or ignored, the (possibly) resulting personality disorder is more difficult to treat or interrupt (Thapar et al., 2015).

1.2.2 Prevalence

The most comprehensive method of accessing rates for mental health and behavioural disorders in children is through the Center for Disease Control and prevention (CDC), a federal organisation operating under the Department of Health and Human Services. The CDC describes these disorders as "serious changes in the way children typically learn, behave, or handle their emotions, causing distress and problems getting through the day" (www.cdc.gov). There are a number of methods for determining which children have a mental health disorder or diagnosis; the CDC uses surveys in which parents report on any diagnosis that their child may have received from a healthcare provider.

In the United States behavioural problems in youth have an estimated prevalence of as high as 20% and are often predictors of later symptoms including depression, alcohol and substance use, lower employment rates, delinquency, removal from the home, conduct disorder and criminal actions (Trentacosta, Hyde, Goodlett, & Shaw, 2013).

By most recent available statistics, the CDC reported that 7.4% (approximately 4.7 million) of children age 3 to 17 are diagnosed with Behavioural Problems, with 9.4% (approximately 6.1 million) children age 2-17 years old having received an ADHD diagnosis within a year; 4.7% of children in the US have been diagnosed with illicit drug use disorder in the past year (http://www.cdc.gov). These issues have been shown to be more likely in boys than girls. It may be that girls are less likely to externally manifest their symptoms, or boys come to the attention of teachers and authority figures more easily and frequently (Thapar et al, 2015). It is known that there is a great deal of under reporting and minimizing that occurs in this area, and it is noted by the CDC that this percentage does not apply to the entire previous lifespan of the child (www.cdc.gov). The World Health Organization (WHO) reports that outside of the US, disorders such as Attention Deficit Hyperactivity Disorder (ADHD), which often may result in behaviours

which are cause for concern particularly in the school setting, has a prevalence of approximately 5% globally. A literature search reveals a disparity in the reporting of prevalence, and the American Psychiatric Association (APA) and WHO report that prevalence of behaviour problems in youth may vary from 2-16% in the global population, with some reviews reporting prevalence as high as 20%, specifically in India and Sub Saharan Africa, with suicide being one of the leading causes of death amongst young people (Gupta et al, 2017). It has been estimated that as much as one-third of children with early childhood behaviour problems go on to develop Conduct Disorder (CD) (APA, 2000; WHO, 1992), which is a far more persistent and antisocial pattern of behaviour, and is a diagnostic link to a later diagnosis of Antisocial Personality Disorder (APD) (DSM-V; Thapar et al., 2015).

Also, children and youth who witness or are exposed to domestic violence or trauma have been shown to have a higher prevalence of behavioural problems (Fergusson, Woodward, & Horwood, 1999).

1.2.3 Aetiology

The distinction between normative and problematic development is often obfuscated by the fact that oppositionality, argumentativeness, emotional lability and some defiance of authority are normal hallmarks of any healthy development. Psychological theoretical models agree upon the fact that these behaviours are, to an extent, present in individuals regardless of any maladaptive behaviour. If and when the disruptive patterns become pervasive and persistent, or when they compromise the healthy and normal psychosocial functioning of the individual, only then are they diagnosable as a behavioural disorder (Campbell, Shaw, & Gilliom, 2000).

Any number of factors have been shown to have an impact on the development of a behavioural disorder, such as genetic predisposition, biological differences, parenting and familial factors, environmental impacts lower socioeconomic status and other individual differences.

Psychopathology of one or both parents has also been shown to have an effect on behavioural disorders, as does parental substance misuse, interpersonal violence and divorce, separation or the death of a parent or caregiver.

Children and youth who have increased difficulties in the school setting, intellectual deficits or underperforming may have a higher incidence of behavioural disorders, and there is a high rate of comorbidity of ADHD and CD or ODD (Thapar et al., 2015).

The ability to process stressors and self soothe may also play a role in the development of behavioural problems as well, this ability is known as social cognition, and when it is disruptive or nonlinear in nature, there may be a higher likelihood of behavioural disorders. Children and youth who are more aggressive have a tendency to respond to situations and events in a more reactive manner, often perceiving threat or danger in situations when there is none (Kempes et al., 2005). Because they perceive situations as being a threat, they may in turn learn to rely increasingly on their maladaptive behavioural patterns and exhibit these problems with increasing severity. In short, when an individual is not able to cope with frustration or tolerate difficult emotional and interpersonal situations, he or she may not be able to behave in a socially acceptable or diagnostically healthy manner.

1.2.4 Childhood/Familial Risk Factors

Childhood developmental theories have widely acknowledged and shown that childhood development and parental styles, actions, responses and interactions all contribute to the manner children learn to behave and how that behaviour manifests later in life.

Bowlby's theory of attachment posits that a consistent and healthy attachment to at least one primary care giver is central to an individual's becoming psychologically balanced and being able to form secure bonds and relationships (Bowlby, 1969). Lessons imprinted early on in development have been determined to be in part predictive and contributory to actions and conduct.

Attachment Theory is perhaps the most widely accepted and generalizable model, suggesting that the manner in which parents interact and communicate with their children is directly observable in the manner in which they approach the world around them. In Attachment Theory, a child will develop one of three attachment styles: secure, insecure avoidant and insecure ambivalent/resistant (Ainsworth, Blehar, Waters, & Wall, 1978; Bowlby, 1969). When a child does not feel secure, or safe, in the knowledge that a parent is present – both emotionally and practically – he or she will not be able to develop a healthy and secure attachment style. Children who have been raised with neglectful parenting styles, or inversely with parents who are overly attached and demanding, will be far more likely to be inclined towards conduct and behaviour issues (Bowlby, 1969; Thapar, et al., 2015).

The relationship between a parent and child is complex, and several factors may contribute to the child becoming less able to adapt to stressors in a healthy and consistent manner. Parents who have their own psychological or psychosocial issues will be far more likely to negatively impact the behaviour and development of their children. Maternal depression, anxiety, post-partum depression, and mood disorders all have been shown to negatively impact parenting styles. Those individuals who suffer from such disorders have been shown to be more likely to engage in physical punishment, emotional abuse or neglect, as well as a simple failure to be consistent and to provide adequate supervision to their children (Thapar et al., 2015)

1.2.5 Environmental Risk Factors

While familial impact is shown to be vital in the development of a child, it would be remiss to dismiss the importance of the social and cultural environment in which a child is raised.

Bronfenbrenner's development of ecological development places a child within five levels of relatedness: microsystem, which consists of the immediate and closest relationships such as those parents and family; mesosystem, which consists of the relationships one step removed from

the child, such as parental relationship with the child's peers; exo-system, which describes the linkages between external settings not directly related to the child but still effecting him or her; the macrosystem can be described as the culture or communal patterns in which the child is within; the chronosystem refers to the influence of time, and acknowledges the changes and variations resulting from this (Bronfenbrenner, 1979). This theory demands that a child be placed into his or her context, and that it is not just the family but also these larger systems that effect development and behaviour.

Epidemiological and longitudinal research has demonstrated that higher levels of poverty, unemployment, criminality and catastrophic events are consistently demonstrative of the relationship between these multiple stressors and a higher occurrence of mental disorders and behavioural problems (Garbarino, 1990; Sektan et al., 2010). Furthermore, when there are combinations of stressors, such as a low income combined with a recent parental separation or loss, a child becomes vulnerable and his or her developmental is more likely to be adversely influenced (Prochnow & DeFronzo, 1997). While these elements have been repeatedly shown to be contributory to child misconduct and behavioural or emotional problems, it is critical that we remember not all children being raised in so-called adverse environments display any of the aforementioned behavioural issues. The familial and parental interaction remain at the core of a child's development, and a child can, and often does, grow into a healthy and well-adjusted adulthood despite adverse conditions if he or she has been raised in a supportive, consistent and safe family system. (Goldstein & Brooks, 2005; Thapar et al., 2015)

1.2.6 Impacts

Behavioural problems are often predictors of later symptoms including depression, alcohol and substance use, lower employment rates, delinquency, removal from the home, conduct disorder and criminal actions (Barlow, 2012; Thapar et al., 2015). A number of longitudinal studies

containing appropriately matched control groups have shown that conduct and behavioural problems are predictors of a broad range issues in later adolescence and extending into adulthood. These studies show that boys with externalising problems earlier in life have a much higher rate of not just psychological disorders but also adverse life events including unstable friendships, inability to experience healthy intimacy, job terminations, incarceration and a number of other problems with individual and social stability (Bongers, Koot, Van der Ende, & Verhulst, 2008; Ferguson, Horwood, & Ridder, 2005; Robbins, 1996). Coleman et al (2009) found that in a 40-year follow up adolescents with even mild externalising symptoms had negative outcomes on a composite peer reviewed measure rating adverse symptoms in adulthood.

This group of disorders affects not only the youth and family but extends to have a negative impact on communities and society at large. Juvenile delinquency is often a result of these behavioural problems, with home and school based acting out transitioning into criminal offenses (Robbins, 1996; Thapar et al., 2015). Within the family setting, attention and care may be focused on the identified client to the detriment of siblings or the family as a whole. The balance within the family can be disturbed, and as imbalances and disfunction increases or persists there are resulting higher rates of divorce, abandonment, alcohol and substance use, untreated mental health issues and other maladaptive behaviours. Abuse, neglect, interpersonal violence and sexual misconduct are possible results of family dysfunction; evidence has shown this correlation through experimental, longitudinal and theory-based scholarship (Thapar et al., 2015). The impact of adolescent behaviour problems and juvenile delinquency is highlighted by the fact that in the United States a person under the age of 18 commits 1 in 12 domestic violence offenses that come to the attention of law enforcement (http://www.ojjdp.gov/ojstatbb).

In addition to the repercussions to the family, increased violence and delinquency within a community leads to lower property values that directly impacts school funding, growing unemployment rates and increased illegal activities such as substance use, underage alcohol consumption and the sale of narcotics (Petterutti, 2009). These activities may lead to imprisonment that will in turn affect a youth's ability to contribute to the long-term economic productivity of the community and the nation as a whole. Over 800,000 minors age 10 to 17 were arrested in 2016, with approximately 200,000 of these arrests being classified as a violent or weapons related charge. The cost to taxpayers is significant and avoidable; the Justice Policy Institute has estimated the average cost of incarcerating a juvenile to be $148,767 per person per year (justicepolicy.org); all told, the United States spends approximately 5.7 billion dollars per year on the incarceration of minors (http://www.ojjdp.gov/ojstatbb). By comparison, alternative treatment options that include treatment or counselling may cost closer to $15,000 per year (justicepolicy.org;ojjdp.gov).

The United Kingdom reported 88,600 arrests of minors in the year 2016, with the most common offenses being violence against the person (26%), theft offences (14%), and criminal damage (12%) (www.gov.uk). What these numbers indicate is a serious problem in both the US and UK, with large numbers of violent crimes being committed by minors, and significant expenditures being allocated to the treatment and incarceration of minors.

1.2.7 Who Refers Youth for Treatment

Referrals to treatment come from any number of sources, from a concerned parent, to a teacher, to an arresting officer or the family court judge involved in a young person's case. Teachers will often be the referral source, as they might also witness bullying or aggression in the classroom, withdrawal from social or healthy interactions with other students, failure to progress or succeed academically and any number of other warning signs that may require the intervention of authorities or the opinion and diagnosis of a professional therapist or clinician. A referral may

also come from a mandated reporter: a person who, because of his or her job, must contact the appropriate authorities immediately when he or she suspects, witnesses evidence of or is told that a child is being neglected or abused.

1.2.8 Who is Referred for Treatment

It is known that there are more young people who are in need of treatment and intervention than come to the attention of authorities or adults (Thapar et al., 2015; www.ojjdp.gov/ojstatbb).

Referrals to treatment are, broadly, either emotional or behavioural in nature; unfortunately, many children, particularly those with depression, anxiety, or other (largely) internalising diagnoses, may not be referred unless their symptoms are severe enough to draw the attention of others (Thapar et al., 2015). Self harm, suicidal ideation or attempt, panic attacks or similar will likely result in a referral.

Adolescence is a time of biological, environmental, developmental and social change, which leads to increased 'acting out' behaviours, as well as more extreme or dramatic actions by the individual. These are likely externalizing behaviours and it is often the case that such behaviours are indicative not merely of behavioural misconduct but problems in the home or psychological distress of the child. While minority or lower income children are more often referred, and subsequently diagnosed, with behavioural disorders such as ODD or CD, white middle and upper class children are more likely to be referred for ADHD, for which they are frequently prescribed medication and not involved further in child welfare or juvenile justice systems (Morgan et al., 2013). What this suggests is that the same externalizing behaviours, such as impulsivity, academic problems, irritability and aggression, may lead to a white child being referred to treatment for ADHD, and a minority or lower income child being identified as disruptive or defiant, and subsequently becoming involved in more punitive systems. Once a child

is identified and labelled as disruptive, resistant, or otherwise behaviourally disordered, he or she becomes far more likely to be suspended, expelled, or placed on probation for minor infractions.

The school setting produces an increasing number of referrals, in part because children and youth spend the majority of their time in classrooms, with teachers observing them and taking note of any disruptive or troubling behaviours. ADHD, either undiagnosed or unmanaged, will often result in behaviours that garner the attention of authorities. These behaviours include difficulties sitting still, paying attention, trouble completing tasks or remaining focused for sustained periods of time. A young person's grades and achievement in the academic arena may be adversely affected by these issues, which often results in therapeutic or behavioural interventions. Persistent truancy will also alert authorities to possible problems in the home, with the family or other issues relating to emotional or behavioural troubles.

While adolescence contains a healthy amount of sexual development and experimentation, there are certain behaviours which are cause for concern, and may lead to referral and intervention. These behaviours include a pattern of unwanted and uninvited touching or fondling of peers or younger children, a preoccupation with pornography that adversely affects other prosocial activities, contact with older and more developed persons, surreptitious watching or 'peeping' behaviours and more extreme or violent acts of sexual coercion or assault (Ryan, Leversee, & Lane, 2011). In the United States, the age of consent ranges from 16 to 18 years, in the UK the age of consent is 16, and in some countries, such as Burkina Faso, Comoros and Japan, the age of consent is 13. Minors are, technically, prohibited by law from engaging in sexual activity. However, it is rare that these offenses are prosecuted or reported, unless the gap in age is significant and cause for concern (fbi.gov)

It is important to note that sexual behaviour, without the element of maladjustment or harm to oneself or others, is a normative and understandable part of an individual's development, and is not cause for alarm or referral. Teenage pregnancy or diagnosis of a sexually transmitted disease may alert a health care provider to the unsafe sexual activity of a young person, and while a family doctor or general physician may in some cases be prohibited from making a referral due to confidentiality laws, some instances may require, in accordance with governing statutes, a referral or report. It is vital that practitioners are trained to discuss these risks with their patients. These data, when gathered appropriately and consistently, certainly provides valuable information regarding the health and safety of adolescents.

A myriad of behaviours or physical signs may indicate that a child is being abused in the home or elsewhere, and these behaviours will often lead first to a report, an investigation and eventually to a course of treatment or in some cases removal from the home. Physical abuse may be indicated by frequent visits to the hospital or clinic, signs of bruising, marks or exaggerated startle reflex; sexual abuse may be indicated by the adolescent's sexual acting out or misbehaviour such as increased and age-inappropriate promiscuity or overly sexualized behaviours; often victims of sexual abuse may be unreasonably afraid of being alone or the dark, extremely irritable or emotionally labile, there may be the presence of enuresis, hair pulling, cutting, self-harm, substance use, insomnia, nightmares, panic attacks, running away or suicidal gestures or attempts (Thapar et al., 2015). Research has shown that often the behaviours which have brought a child to the attention of authorities are in fact a response to problems in the home or with primary caregivers.

1.3 Approaches to Treatment and Prevention

A number of treatment approaches and interventions have been designed, developed and researched in the ongoing effort to prevent, reduce and treat these damaging and harmful

behaviours. The research base for these interventions is disparate, with some evidence indicating that family based or systemic approaches may be the most effective.

There are a variety of interventions targeting youth age 10 to 17 who are at risk of being removed from the home or have otherwise come to the attention of authorities due to antisocial behaviour or mental health diagnoses. The interventions differ by program theory, intensity, and duration. While in some cases these programmes result in incarceration, the majority do not, and behaviourally disordered youth are often mandated to treatment or out-of-home placements either in foster care, treatment facilities or group homes. The treatments may consist of psychiatric medication, individual treatment or, increasingly, a course of systemic therapy as both a preventative and rehabilitative measure (Petterutti et al., 2009). There are a number of guidelines available that contain within them recommendations for evidence-based or research supported complex interventions. The Blueprints for Violence Prevention, (supported by the Office of Juvenile Justice and Delinquency Prevention), the World Health Organization, and the National Institute of Health, to name but a few, all propose guidelines, treatment options, and behavioural, mental health and systemic approaches (blueprintsprograms.org; who.int; nih.gov).

Of course treatment, or prevention, can also be framed in more punitive terms, and these measures range from a course of what is defined as 'simple' probation - consisting of cursory check-ins and perhaps an ankle monitor - to sentencing in a secure detention facility with a number of options between these two extremes.

1.3.1 Pharmacological Interventions

A recent review examined possible benefits of pharmacological treatments for aggression and behavioural problems. The review included 10 trials containing a total of 896 children and youths aged 5 to 18 years. (Loy, Merry, Hetrick, & Stasiak, 2012). The authors conclude that due to the fact we do not yet understand the aetiology of aggression or conduct disorders it is difficult

to target underlying neurological or chemical issues. Anti-social and behavioural qualities are extremely complex and disparate; attempts to remedy them with psychotropic medications have been unsuccessful and will likely continue to be. A Cochrane systematic review suggests that, apart from the immediate tranquilizing effects of the antipsychotic risperidone, which may result in an alleviation of aggressive, impulsive or disruptive behaviours, there is little evidence that drug treatment is appropriate in this context (Pappadopulos et al, 2006). These sedative effects are not lasting and will not be sustained once the medication is discontinued (Loy et al., 2012).

1.3.2 Individual Psychodynamic Approaches

With the possible exception of the more structured and outcome focused approaches, traditional psychotherapy has largely been shown to be an only moderately effective intervention technique for the reduction of distress or externalizing behaviours in youth. The empirical evidentiary base relating to psychodynamic treatment is quite limited, in part due to the possible 'vagueness' of treatment goals, the uncertain and often unnecessarily extended duration of treatment, the unavailability of standardized treatment outcomes and the possibility of reliance not on empirical data but on case reports, individual studies and expert or practitioner opinion (Thapar et al, 2015). In a trial conducted by Fonagy and Target (1994), 135 adolescents diagnosed with disruptive behaviour disorders were matched with others diagnosed with emotional but largely internalising issues. The trial indicated that those with emotional diagnoses improved at a significantly higher rate than their behaviourally disordered counterparts.

Notably, individual psychotherapy is becoming increasingly rare as a treatment approach for adolescents with externalising symptoms, especially if it is not partnered with parent and/or family work. In a retrospective study conducted by the Anna Freud Centre, it was shown that pairing individual with family and parent work is more likely to result in a positive outcome for the child (Fonagy & Target, 1994).

1.3.3 Cognitive Behavioural Approaches

Some research has shown that a potentially effective treatment, particularly for youth facing traumatic events, is Cognitive Behavioural Therapy. CBT is an insight-oriented treatment used in the management of and coping with stressors or negative emotions and feelings. Cognitive therapy was originally developed in the late 1960's by Aaron Beck and is widely acknowledged by mental health professionals as a preferred mode of treatment (Beck, 1967; Thapar et al, 2015). Cognitive approaches to treatment focus on helping the patient become aware of negative or harmful interpretations of their emotions and aiding people in reducing these ways of thinking which will in turn lessen psychological symptoms and alter behaviour in positive ways (Beck, 1967). Because CBT is time-limited, it has been favoured as a treatment option for many different populations. Some suggest that CBT, often in addition to medication, can be beneficial in the lessening of externalizing behaviours in adolescents (Lochman et al., 2011; Thapar et al., 2015). When examining recidivism rates for youthful offenders, CBT has been shown to have some positive effects, with the results of multiple reviews in favour of CBT compared to standard treatment or no treatment, with a reduction in recidivism of approximately 10% on average (Armelius & Andreassen, 2007; Landenburger & Lipsey, 2005). In a (2009) meta-analysis, Lipsey reported that the therapeutic approach of CBT was more effective in the reduction of further criminality in youth than the other included interventions; However, the positive effects of CBT may not be sustainable over time (Johnson & Friborg, 2015). Armelius and Andreassen (2007) found no significant or lasting effects of CBT for youth after 24 months.

1.3.4 Residential Treatment

A national survey conducted in the United States stated that of youth aged 13 to 18 years, 30% were undergoing treatment in a residential facility (Pottick, Hansell, Gaboda et al., 1993). These numbers have significantly decreased since that time, but inpatient units continue to be a

treatment option for severely disturbed or psychiatrically disordered youth. Youth will not likely be referred to residential treatment or committed to a secure treatment facility unless they are in significant danger and pose a risk either to themselves or to others. It is only when a young person's behaviour has become so difficult to control, so dangerous or unmanageable that he or she cannot maintain a necessary level of safety and functioning in the world that he or she will be referred to a psychiatric treatment facility or locked ward. Behavioural problems such as extreme violence, sexual crimes, psychosis or recurrent, persistent and severe suicidal or homicidal ideation, intent or attempts will often lead to inpatient treatment for a proscribed or court mandated period of time (Thapar et al., 2015). Emergency admissions to a psychiatric unit or hospital are required by law in certain situations, and therapists are required to have youth admitted based on the presence of overwhelming risk factors and the inability of a child to function or remain safe. In the United States, the Health Insurance Portability and Accountability Act (HIPAA) of 1996 requires that confidentiality be broken in cases of severe risk. These guidelines are also set forth in the ethical guidelines governing psychiatry, social work, psychology and other counselling professions. Similarly, the Mental Health Act of 1983 is the United Kingdom's primary legislation relating to the treatment and rights of those with mental health diagnoses.

Any effective residential treatment will stabilize the individual and plan extensively for release, it has been shown that a failure to foresee or adequately prepare for discharge will result in an increased likelihood for recommitment (Pfeiffer & Strzelecki, 1990).

1.3.5 Juvenile Justice Interventions

In both the United States and the United Kingdom, legal interventions include probation, electronic monitoring, group home placement and juvenile penitentiaries. While the number of first time offenders has decreased, there continues to be a high rate of reoffending for youth in the court systems (www.ojjdp.gov/ojstatbb). The 2016 Youth Outcome Report in the United States

reports that while in custody, youth often do not undergo treatment of any sort (www.ojjdp.gov/ojstatbb); a review published in 2012 states that failure to be involved in treatment post sentencing leads to an increased likelihood of recidivism and longer-term removal from the home (Newman et al., 2012). Juvenile justice interventions consist of a number of different approaches, but all of them can be described as punitive and involving less treatment than research supported alternatives.

Outdoor Behavioral Healthcare (OBH), or what is more commonly known as wilderness therapy, are privately owned facilities and thus entirely unregulated, but very often receive their referrals directly from the court system. In these programmes, youth are usually housed together in dormitory-like settings, or sent on days or even weeks long outdoor hikes where they are expected to learn wilderness survival skills such as building shelter, foraging for food, and fire starting. These programmes differ in length and intensity but have come under increased scrutiny for cases of abuse, neglect, sexual assaults and indentured servitude. There is no evidentiary support for these programmes.

Similarly, juvenile courts may sentence a youth to a boot-camp style facility, which are overly harsh and in some cases even abusive. A course of Scared Straight, a programme in which at-risk youth and juvenile offenders are exposed to first-hand observation of prisoners and prison in the attempt to deter them from future or continued criminal behaviour has been demonstrated to cause significant harms (Petrosino, Turpin-Petrosino, Hollis-Peel, & Lavenberg, 2013). These programmes have been shown not only as ineffective, but also to increase the likelihood of further and more severe criminal activity.

The Cambridge-Somerville study, a longitudinal study commissioned in 1936, randomised 506 boys age 5 to 13 years into either a treatment group or control group; the boys assigned to the

treatment group received psychiatric treatment, medical treatment, mentoring and psychological counselling in an effort to prevent juvenile delinquency and criminal behaviours. Decades after the original ten-year follow up, which showed that there were no effects of the intervention, McCord published work demonstrating that there were in fact harmful effects on those in the treatment group (McCord, 1978). Cambridge-Somerville has become synonymous, and shorthand, for the presence of stigmatising and peer effects. In programmes in which youth are placed together with their misbehaving peers, they are likely to bond with one another and encourage deviant actions and attitudes (Bandura & Walters, 1959).

Despite overwhelming evidence showing that such interventions do not result in positive outcomes, overly harsh, boot-camp style or purely punitive measures and programmes continue to be used in the United States, often with the financial backing of the state; these programmes have been demonstrated by multiple reviews to have either no impact or a negative result for participants (Petrosino, Turpin-Petrosino, Hollis-Peel, & Lavenberg, 2013; Wilson & Lipsey, 2000).

1.3.6 School Interventions

School based interventions to deter delinquency, violence and behavioural misconduct have a significant amount of support, and there are a number of strategies that administrators and stakeholders may choose to adopt, including surveillance such as video cameras, security guards and metal detectors, deterrence measures such as zero-tolerance policies, and psychosocial or interpersonal training programmes (Durlak, 1995).

In a meta-analysis published in 2007, researchers reported data collected from 249 studies, ninety percent of which were conducted in the United States and all of which used disruptive or aggressive behaviours as an outcome measure. The analysis showed that there are a number of interventions available for schools to choose from, all of which have some reported positive

effects. The authors showed that it is predominately through a combination of the available approaches that the strongest effects are achieved (Wilson & Lipsey, 2007).

1.3.7 Parent Based Interventions

Developmental theories agree that the parental interactions between a child and his or her parent(s) are formative in healthy development (Thapar et al., 2015; Behan & Carr, 2000). It is therefore reasonable to assume that working with parents on their parenting skills and style may result in positive outcomes for a behaviourally disordered child. Of these interventions, most are strength based and focus on engagement with one's child and positive interactions and well as boundary-setting, reasonable supervision and understanding. Parent training is often delivered in a group setting, with differences between programmes but usually consisting of meetings and weekly homework. The more researched of these are manualised and facilitated by a specifically trained therapist. Two such programs that are both supported by best practice lists are: 1) Incredible Years, (Webster-Stratton, 2006) a programme developed by Caroline Webster-Stratton and delivered in over 24 countries, which is a brief (12-20 weekly sessions lasting 2-3 hours) group parenting intervention that focuses on improving parenting skills and promoting engagement with their children's school activities and learning and 2) Positive Parenting Program (known as Triple P), (Sanders et al., 2014) which is also designed to aid parents in setting goals, reducing dysfunctional parenting practices and thereby reducing risk factors for children and families (blueprintsprograms.org)

1.3.8 Systemic Approaches to Family Treatment

Systemic approaches to family treatment have their foundations in a long history of theory and practice. Until the late 1960s and 70's, most clinicians and practitioners viewed youth and juvenile problems as extensions of individual psychopathology. Clinicians were for the most part still under the influence of Freudian or Jungian theory, and approached analysis and treatment

accordingly. In this model, behaviour is seen as a manifestation of an entirely internal and emotional state, and approached thus (Freud, 1955). Because an individual cannot be treated without consideration of his or her environment, and with an increasing interest in the treatment of adolescents, focus began to shift towards a more holistic and environmental attitude. Erikson's theory of cognitive development, which presents 8 distinct stages in the identity and psychosocial development of an individual and considers the impact of external and cultural factors on this development (Erikson, 1956), emerged and began the shift away from the Freudian. Similarly, Bandura's theories had a strong impact when he suggested in the early 1970's that there should perhaps be more emphasis placed on the relational patterns and influences between people and families (Bandura, 1977).

These theoretical shifts in combination with Bronfenbrenner's theory of ecological development described in more detail in section 1.2.5 of this thesis gave rise to structural family therapy, a conceptual framework and treatment approach designed by Salvador Minuchin. Structural Family Therapy seeks to understand the form and structure of a family, and for the therapist to join with the family thus becoming a motivator for change. Minuchin developed a method of mapping and portraying the structure and relationships within a family. He sought to interrupt negative patterns and believed that dysfunction lay with the family as an entity, not with one particular identified individual within it (Minuchin, 1974).

Out of this emerged Strategic Family Therapy, an expansion and interpretation founded by Jay Haley and further developed at the Family Therapy Institute in Washington, DC. Haley's theory is far more directive in nature and consists of planned and carefully designed interventions (Haley, 1976). Strategic therapy also views all behaviour, even negative or destructive, as attempts

to communicate. Actions are intended to convey meaning; symptoms are in fact (often frustrated) efforts to communicate (Haley, 1976).

Systems Family Therapy contains within it the aforementioned schools of thought and theoretical approaches. Both Minuchin and Haley subscribed to, and were influenced greatly, by systems theory, which posits that a child exists in a series of interconnected systems, and any systemic approach to treatment will of necessity engage not only with the child, or even with the child and her family, but also with the larger systems such as the school, peers, church, community, probation and parole, child services and other community elements as needed. The theory suggests that, given the connectedness of all these systems, influencing one while ignoring any of the others will undercut and lessen the value of any intervention or positive change.

When considering that a child's development and growth often begin and are rooted in the familial relationship, it follows that when treating behavioural difficulties the family may be at the centre of the treatment approach. Systems theory and family approaches to intervention are closely linked and one can be seen as informing one another; the beginnings of familial approaches are found rooted in systems theory. The current research base suggests with caution that family interventions may be beneficial to youth aged 10 to 18 presenting with behavioural disorders as well as for those in danger of removal from the home for protective or legal reasons. Family interventions may also reduce the length of time spent by youth in institutions or in care (Woolfenden, Williams, & Peat, 2001). A number of family-based interventions have been developed to prevent and/or treat behavioural problems among children and youth.

Multi-Systemic Therapy (MST) is one systemic approach that has gained favour in some court systems as a mandated treatment for offending or high-risk youth and their families (Henggeler, Schoenwald, Borduin, Rowland & Cunningham, 2009). MST is a home and

community-based treatment approach that is result focused and attempts to reduce out of home placements and delinquency by providing a type of wrap around service that will not only offer home based treatment but also interactions with the school and community of the identified child. MST was developed with elements drawn from strategic family therapy, structural family therapy and cognitive behavioural therapy. Evidence for MST is inconsistent and a recent study shows that multisystemic therapy does not significantly impact outcomes for moderate to severe antisocial behaviour in youth age 10 to 17 (Fonagy et al, 2018). While MST has long been included on best-practice lists and is regarded by many as both evidence-based and efficacious for a broad variety of client populations, this may not be the case. In actuality, a systematic review demonstrated that the intervention has no significant effects and there may be significant allegiance effects present in the reporting (Littell, Popa, & Forsythe, 2005).

1.4 Functional Family Therapy (FFT)

Functional Family Therapy (FFT) situates and aligns itself with this tradition of systems theory. According to the designers of the model, the different theoretical approaches present from the late 1960's to early 1980's were "synthesized and made compatible via the phase-based strategy of FFT" (Alexander, 2013, p. 25). The model, theory of change and explanation of its design is detailed in the subsequent section.

1.4.1 Principles of Functional Family Therapy (FFT)

FFT reflects a core set of theoretical principles, in which behaviour is seen as a representation of the family relational system; i.e., as indicative of the communication, patterns and purposes of the family. Its developers describe the overarching goals of FFT as follows: "To change the maladaptive behaviours of youth and families, especially those who at the outset may not be motivated or may not believe they can change; to reduce the personal, societal, and economic consequences of disruptive behaviour disorders and to do so with less cost, in terms of

time and money, than many other treatments currently available" (Alexander, 2002, p. 2007). FFT uses a manualised approach and includes reframing, interrupting of negativity or blame, redirection of focus, interpretations of patterns of maladaptive behaviour with links to emotions, a deepening understanding of actions, and communication training with focus on positive communication. It incorporates theories of information processing, social cognition, and the psychology of emotion.

FFT is a short-term (90-day), intensive and comprehensive program designed for the treatment of behavioural misconduct in youth aged 10 - 18. It is a treatment modality rooted in family systems theory and cognitive behavioural theory that can be delivered in clinical settings, school settings, or in the home. It is intended to address those youth and families with a wide array of presenting problems including but not limited to; criminal behaviour, truancy, running away, sexual misconduct, substance abuse, risk of out of home placement and as a re-entry program for youth returning to the community following release from institutional settings. The program requires between 8 and 30 hours of direct service to youth and their families typically over an average of 12 home visits in 90 days. Over more than three decades its designers have modified the original concept in efforts to increase effectiveness and uptake by service providers. According to the literature presented by FFT Inc., these changes have been made based on a solid foundation of evidence (Alexander 2013).

Therapists are expected to adjust FFT to family members' capacities and to the specific problems they face. In line with its theoretical underpinnings, the focus of an FFT therapist should not be on one specific member of the family, rather the therapists attempt to build an equal and balanced alliance with each member of the group. This technique is designed and intended to minimize blame, and to foster an environment in which each family member feels understood and

vital to the process. FFT uses reframing, interrupting of negativity or blame, redirection of focus, interpretations of patterns of maladaptive behaviour with links to emotions, deepening understanding of actions, and communication training with focus on positive communication. It incorporates theories of information processing, social cognition, and the structural relationships within families (Alexander 2013).

1.4.2 Core Components of FFT

The intervention consists of five major components and may include pre-treatment activities: engagement in change; motivation to change; relational and interpersonal assessment; behaviour change and generalization across behavioural domains and multiple systems. These goals are accomplished in five ordered phases each of which is dependent upon the phase prior; each phase has an assessment and intervention component directed at specific goals. The phases have been conflated into three (Engagement/Motivation, Behaviour Change and Generalization) until a recent expansion of detail within the FFT literature and training manuals. Details of the five phases are, as explicated by the originators of FFT, as follows:

Phase 1: Engagement, wherein workers focus on the engagement of the entire family, which involves 'maximizing factors which enhance the perception that positive change might occur (intervention credibility), and minimizing factors (e.g., poor program image, difficult location, insensitive referral) that might signify insensitivity and/or inappropriate resources' (Alexander 2013, p. 11). This phase also requires a high level of engagement from the therapist including availability, contact with as many family members as possible, cultural sensitivity and ‘matching’ to the family unit. Therapists must maintain strength based relational focus in this, as well as all subsequent, phases (Weisman & Montgomery, 2018).

Phase 2: Motivation attempts to minimize hopelessness and to create a positive context. ‘To develop or enhance family members’ motivation, workers identify and quickly begin to

modify the pattern of changeable intra family risk factors, especially negativity, hopelessness and blaming; [and] initiate and/or strengthen intra-familial protective factors that can mitigate the effect of risk factors that cannot be changed' (Alexander 2013, p. 12). Therapists focus on the relationships and interrupt blaming sequences while encouraging positive themes and reframing negative behaviours. Reframing, a central theme and principle of FFT, is a three-step process in which the therapist must first identify and make clear the negative aspects of the behaviour or pattern, offer the family a possible alternative meaning which is 'noble' or benign rather than purely negative, and finally go on to elaborate or refine the new meaning, or (if the family rejects the new meaning entirely) be able to move onwards without losing alliance with them.

Phase 3: Relational Assessments are performed in order to understand and analyse information relating to the relational process of the family, as well as to plan for the upcoming behaviour change and generalization phases. Therapists continue to redirect the focus from individual problems to a relational perspective and suggest meaning relating to negative behaviours (Weisman & Montgomery, 2019) .

Phase 4: Behaviour Change is used to "develop long term behaviour change patterns that are culturally appropriate, context sensitive, and individualized to the unique characteristics of each family member"(Alexander 2013, p. 13). Workers focus on cognitive, interactive, and emotional issues; emphasize positive communication and parenting skills; and provide concrete resources that "guide and symbolize specific changes in behaviour" (Alexander 2013, p. 13). This phase aims to reduce intrafamilial risk factors and enhance intrafamilial protective factors. In this, the most directive phase of treatment, therapists may assign tasks, use technical aids, teaching, role playing, training, conflict resolution skills or other means for developing necessary skills.

Phase 5: Generalization is the final goal of treatment and focuses on the generalization of behaviour change to other settings and social systems. This involves referrals, mobilizing community support systems and modifying deteriorated family-system relationships with schools, probation officers or other systems. “Generalization activities involve knowing the community, developing and maintaining contacts, initiating clinical linkages, creating relapse prevention plans and helping the family develop independence” (Alexander 2013, p. 14).

After the conclusion of treatment, families are often seen in Booster Sessions; these sessions can be requested by the family, or as a refresher if there is a relapse of maladaptive behaviours. A relapse may be signified by objective outcomes such as arrest or truancy, or by family reporting or desire for assistance. These sessions are designed to reinforce the skills learned in the Behaviour Change Phase of treatment, and to remind the family of their ability to achieve a higher level of functioning.

Underpinning the five phases is the concept of ‘obtainable change’; this refers to the setting of realistic and achievable goals. These goals are not merely those set by the therapist, authorities, probation officers or other treatment providers, rather agreed upon and discussed within the scope of treatment. The youth and family are integral to the decision of what goals to pursue in treatment (Weisman & Montgomery, 2018).

Assessment focuses on the functional nature of problems within the family rather than on a diagnosis, and is described as a “continuous, multilevel, multidimensional, and multimethod process that includes individual, family, behavioural, and contextual factors” (Alexander 2013, p. 22). The relational assessment examines not only the connectedness or relative autonomy betwixt family members, but also the hierarchical standing of each member in relation to the others. It focuses on why or how the family continues in its promotion and maintenance of problematic

sequences, chains of behaviours, events and interactions; assessment will also identify key risk and protective factors for the family, primarily to increase the positive whilst decreasing the negative.

1.4.3 Theory of Change

The model is suggested to be useful for complex and multidimensional problems because of its flexible structure and cultural sensitivity. The attention to relational patterns and clear focus on family systems is often stated to be at the core of the intervention's success. As many FFT therapists provide services in the home, it is particularly suited to those clients who may be initially unwilling, unable or otherwise unlikely to attend sessions in a clinical setting. Within FFT there is a very strong stance against blaming and negative treatment or approach. Neither individuals nor families are to be seen or treated as 'bad' or wrong. Often the families treated by an FFT therapist may be described as resistant to treatment, but FFT attempts to overcome this through its use of the techniques discussed above. When there is a perceived failure or lack of improvement, it is the therapist who must identify ways to alter her approach and treatment of the family, not vice versa. FFT therapists attempt to provide a therapeutic experience unlike what most families are accustomed to: an experience that focuses on an alliance and fostering a cooperative working relationship approach. Families should feel engaged and aware rather than belittled or blamed for their problems. The founders and advocates of FFT attribute its effectiveness to the careful sequencing of techniques, helped by the continuous assessment and intervention processes (Alexander 2013; Weisman & Montgomery, 2018).

1.5 Aims of the Research

The purpose of this work is to first clarify the evidence base supporting FFT in a rigorous manner, and an overview of the current evidence must therefore be the starting point for this project. The feasibility and acceptability of FFT must also be ascertained across varied settings

and with heterogenous populations, so an implementation analysis of current FFT studies focuses on explicating any issues and disparities emerging from the overview.

1.6 Summary

The aim of this work was to assess the evidence for the effectiveness and any harms of the FFT programme, using an overview of current evidence and an implementation analysis to explore variations in the literature found. It is with these two pieces of work that FFT can be better understood in its current context. The research supporting FFT will therefore be better understood and decisions regarding its use based in a more reliable and systematic evidentiary analysis than was previously available.

Within social work, a number of treatment options are available to practitioners and agencies, and the decision to use one over another should be made with the best possible evidence and supporting information. This thesis contributes a valid and dependable source for this information and can provide a comprehensive and manageable tool for making treatment decisions around the treatment of youth presenting with behavioural problems.

Current and long-standing evidence suggests that youthful behaviour problems are indelibly linked to environmental factors, the most influential of which may be the youth's family. Without effectively addressing these familial patterns and relationships, a child is more likely to continue in maladaptive behaviours that will become increasingly problematic and harmful, not just to the individual but to the family, community, and society at large. Family based approaches to treatment have been in use for decades but are continuing to rise in uptake. Evidence or research based interventions are being put forth as the most effective means of addressing and treating childhood behavioural problems. Among these treatments, FFT continues to grow in its dissemination and uptake with it being delivered in over 300 sites worldwide (Alexander, 2013;

blueprintsprograms.org). Despite its ever-growing use, FFT does not have a consistent or rigorous evidence base that supports all of the claims posited by the literature surrounding its use.

Continued use of FFT should be examined, and the feasibility and acceptability of FFT must be ascertained across varied settings and with heterogenous populations. The purpose of this work is to first clarify the evidence base supporting FFT in a rigorous manner, and an overview of the current evidence must therefore be the starting point for this project. Following this work, an implementation analysis of current FFT studies will focus and explicate the issues and disparities emerging from the overview. It is with these two pieces of work that FFT can be better understood in its current context. The research supporting FFT will therefore be better understood and decisions regarding its use based in a more reliable and systematic evidentiary analysis than was previously available.

2 METHODOLOGY

2.1 Outline of Chapter

This chapter will lay out the rationale and methods for this thesis; first, the decision to conduct an overview of reviews, or umbrella review, of Functional Family Therapy (FFT), and then to go on with a secondary implementation analysis. First, it is important to understand the concept of bias, and where overviews fit into the range of potential research methods for the questions at hand, as well as why this approach was deemed most appropriate and useful in this case; so, to this end, the hierarchy of evidence will be outlined, as well as where overviews fit into the hierarchy. Importantly, a broader approach will also be discussed, as a more nuanced understanding of how to best answer the specific research questions. Second, it will be explained why the evidence from the overview led to an implementation analysis, as a picture of evidence emerged suggesting that implementation elements and concerns seem to affect the core outcomes of FFT. The evidence supporting FFT is so variable, reporting on different outcome measures, follow-up times, utilising different methodologies and covering such a broad range of outcomes and results that the focus provided by the overview helped a great deal in formulating the approach taken. Thirdly, implementation and fidelity as broader concepts will be explored and contextualised; there are a number of differing and at times conflicting positions within implementation science and choosing a standpoint to approach the implementation of FFT is placed within the broader field. Finally, the Oxford Implementation Index will be presented and the reasoning behind its use explained.

2.2 Bias

Bias is an error, intentional or otherwise, that results in an inaccurate or untruthful portrayal of the results reported. They can lean in either direction, and can be considered as minor and having little impact, or as being significant enough to lead to the entirety of the outcomes being attributable to some form of bias (Higgins, 2011)

Research has conclusively demonstrated that conflict of interest and bias in any form can dramatically influence the outcomes and reporting of intervention trials and reviews (Shadish, 2002). It is therefore crucial to outline what these possible forms of bias may be, and how such biases may impact outcomes. Sources of bias in clinical trials are specifically addressed by Cochrane Handbook guidelines, and include selection bias, performance bias, attrition bias, detection bias, reporting bias. Notably, there are also 'other forms of bias', which include fraud, deviation from protocol, and perhaps most germane in this case, allegiance bias.

2.2.1 Selection bias

This refers to the differences between the individual characteristics of the people or groups being compared. It is through randomization that this form of bias will be minimized. In order to adequately assess whether this was the case in a trial, the processes for allocation concealment must be reported.

2.2.2 Performance bias

This occurs when one group in an experiment receives more attention than another, in other words, all participants should be exposed to similar conditions. However, in the case of complex interventions in which groups are receiving vastly different psychological or social interventions, this is less of an issue.

2.2.3 Detection bias

This relates to how outcomes are determined, and is also aided greatly by blinding of the assessors. This is especially important when the outcomes being measured are subjective, such as self-reported family conflict level.

2.2.4 Attrition bias

This refers to the differences between those who withdraw from the study, or are excluded, which is the case when participants are not included in the reports of analysis, even

when the data is available, as opposed to when outcome data is not available due to the participant leaving the study or being unavailable for follow up or outcome measurement.

2.2.5 Reporting bias

This refers to those findings which are included in the reporting or not. Reporting bias contains within it a number of elements. The first of these, publication bias, or 'the file drawer problem' suggests that studies reporting significant or positive effects are more commonly published than those which report null or negative findings (Higgins, 2011). Reporting bias may also refer to outcome reporting bias, in which a number of outcomes may be noted and measured, but not all of these are reported (selective outcome reporting). The decisions regarding what to report (or not) may lead to published reports being incomplete pictures of the evidence or directly and purposefully misleading.

2.2.6 Conflict of interest and allegiance bias

When examining conflict of interest, it necessary to expand upon the narrow definition which only addresses whether there is a direct financial interest in the research outcomes (Higgins, 2011). This type of conflict is meant to be reported and acknowledged by trialists and reviewing authors. However, this financial conflict should not be limited but is applicable when even one member of a team has an interest in the outcomes being reported or presented (Eisner, 2009). It should be extended to include not just developers but other licence holders and persons for whom future success of the program is central to their career advancement or employment (Eisner, 2009). The promotion of a particular program may also be influenced by developers and researchers attempting to have their programmes included and promoted by best practice lists such as Blueprints for Violence Prevention or World Health Organization guidelines for practice.

Allegiance bias has often been narrowly defined, as a conflict of interest stemming from a strictly financial interest in the outcome. However, more recent literature highlights the fact

that conflict of interest should be determined far more broadly and can include such things as a researcher's need to defend an ideological position, a greater likelihood of publication, career advancement, prestige, and inclusion of the intervention of 'best practice' lists (Eisner, 2009).

The role of developers in interventions has come under increased scrutiny, and trials which are led by developers as evaluators show larger and more positive effects than those which are truly independent (Eisner, 2009; Shadish, 2002). The same is true of reviews, in that those reviews authored by individuals who have an interest in the outcome often present a far more positive picture of the intervention. It is this form of allegiance bias, specifically, which can be reduced by independent, scientifically rigorous, and transparent methods.

Allegiance bias should capture all the relevant forms of allegiance that may manifest in the reporting of trials and by reviewing authors. The same bias which may lead to a trialist not including unfavourable outcomes in their published results may also lead to a reviewer failing to include evidence that does not support their position in a review.

There exist two means by which conflict of interest or allegiance bias may result in biased reporting and outcomes. The first of these is intentional, a conscious decision to manipulate and in some extreme cases actively defraud (Eisner, 2009). The second of these is identified as an unconscious or cognitive bias; in this instance, the researcher does not recognize his or her own cognitive or 'self-serving bias'. Psychological research has pointed out that people often have difficulties ignoring their own self-interest in favour of disinterested and impartial evaluation (Moore, Tetlock, Tanlu, & Bazerman, 2006). In short, people tend to favour positions and outcomes that support or confirm what they already believe to be true. This may occur even when they believe themselves to be acting impartially and fairly.

2.3 Reducing Bias: The Hierarchy of Evidence

To decide whether or not an intervention is effective presents a number of issues. Whether the decision regarding its use is being made by a clinician, a stakeholder, policy-maker or agency director, it is vital that the best evidence and information relating to the intervention be available, accessible, reliable, and free from bias. The research supporting the intervention must be relevant, appropriate to the problem and population being addressed, and valuable to the decision-making process (Gibbs & Gambrill, 2002). The research base supporting an intervention may well exist, but being able to place the research in the correct context, with an awareness of how much importance and weight ought to be given to each component of that research, is a task which in some cases may be greatly aided by the use and understanding of the hierarchy of evidence as originally developed by Sackett and colleagues (Sackett et al, 1996); Sackett's hierarchy places different types of research or evidence on a scale, with the evidence at the top of the scale or pyramid being considered to contain the most internal validity and being the 'gold standard' or 'best evidence' for the research question or area (Higgins, 2011).

Figure 1. Hierarchy of evidence

The hierarchy was intended to reduce bias, and to rank evidence in accordance with this goal. What is often considered 'best evidence' – RCTs, systematic reviews and meta-analyses – is on the top of the pyramid as it is believed in many cases to greatly reduce the possibility of bias through the removal of confounding factors. RCTs and systematic reviews of such RCTs are thought to have a low risk of bias due to randomization and are designed to have a smaller risk of systematic errors; the methods employed in a high quality randomized trial will ensure that participants are equally likely to be assigned to one condition or trial group as another, that there is adequate blinding where possible, and that baseline characteristics are controlled for. Primarily experimental methods are designed, through these processes, to reduce selection and performance bias; it is through these methods that the uses of randomization makes it more possible to identify the causality and effect of a specific intervention. Systematic reviews go one step beyond an individual RCT by gathering data from multiple trials and in essence increase the

sample effect size. Expert opinion occupies the lowest level of the hierarchy, as it is essentially based upon the experiences of an individual, and will by definition have a high likelihood of bias. Moving up the pyramid, as evidence progresses from individual examples to large scale randomization and systematic reviews, the quality of evidence is thought to increase in value and reliability; and while this may in fact reduce bias, it also loses a great deal of detail and depth that may be useful in decision making and understanding of an intervention (Petticrew & Roberts, 2003).

Stakeholders and decision makers must ask the question not just if something works, but for whom? Under what circumstances or contexts? It is possible to gain these details and different sorts of information from other sources, not *instead* of RCTs necessarily, but in *addition* to trials. This is the purpose of process evaluations, user reports, qualitative research and other non-experimental means of research. Speaking to these concerns, Webb (2001) states that "rooting decision making within a behavioral mind-set, evidence-based practice thus ignores the complexity of actual decision making processes in social work" (p. 63). This is a succinct reminder to engage different kinds of research methodologies, to move up and down the pyramid as needed, eschewing the premise it must be adhered to as an intractable hierarchy. Not every approach, no matter how well supported or proven, will be the best or most useful treatment for every family or in every home. Ideally, decisions regarding treatment and care will be based on the evidence, expert opinion/judgment, and the client's own wishes. In rigorous and mindful evidence-based practice, these differences will be planned for, and things tweaked slightly or even dramatically to fit the new environs (Gibbs and Gambrill, 2002). There is no need for the evidentiary portion of this triad to be populated only with RCTs or systematic reviews.

Despite being widely adopted, the hierarchy is considered to be problematic in some cases, and there is an ongoing debate focused on whether it is always correct to determine quality and credibility of evidence based on study design. Randomised controlled trials (RCTs) and systematic reviews or meta-analyses containing *only* RCTs are placed at the top of the hierarchical pyramid, which is contentious for two main reasons; first, because RCTs always occupy the top rung, observational (or other) studies that may in fact be of higher quality than their experimental counterparts will automatically be given less weight; second, the hierarchy does not allow for any methodological flexibility, Sackett and Wennberg (1997) point out that when there is refusal to pay heed to what question is in fact being asked, essentially authors are reduced to "squabbling over the "best" method"; it is vital to remember that different questions are best answered through different types of evidence and methodology (Greenhalgh et al., 2018; Petticrew & Roberts, 2003). Literature has now come to support the stance that methodology may, and perhaps *should* differ from one question to the next, dependent upon the research question being asked. Simply put by Sackett and Wennberg (1997): "the question being asked determines the appropriate research architecture, strategy, and tactics to be used—not tradition, authority, experts, paradigms, or schools of thought" (p. 1636).

The structure of the hierarchy can also be considered too rigid, in that it does not necessarily acknowledge the relationship between different types of evidence, and that many types of research may used to inform each other. For example, objective outcome data may best be gleaned from RCTs, but implementation quality and contextual information will often come from other sources. Hopefully it becomes "less of a choice between extremes than the hierarchy implies, and effective implementation of an intervention ideally requires both sorts of information" (Sackett & Wennberg, 1997, p. 1636). This will add to the depth of understanding

and allow for a more complete picture of the evidence surrounding and supporting an intervention. RCTs are designed to answer effectiveness questions, and these only in the narrowest sense, thus other types of evidence must be used to fill in the gaps, and we can better make determinations relating to specific population, contextual and implementation concerns when we gather and present the different forms of data available.

2.4 Overviews of Reviews in General

2.4.1 Included Study Designs are Reviews

An overview of review, or 'umbrella review' is designed to compile evidence from multiple reviews into one accessible and user-friendly document (Higgins, 2011). An overview is designed to present a comprehensive picture of the existing evidence for a particular intervention, by assessing and examining the existing reviews. In essence, an overview gathers available reviews and uses the *review* as the unit of analysis, while a (systematic or narrative) review uses the *study* as the unit of analysis. In this way, an overview can be seen as being one step removed, or perhaps above, a review. Overview methodology is a newer form of research design, and the parameters and framework regarding their use is not as well defined. Still, Cochrane does lay out guidelines for their completion, which are in many ways very similar to those guidelines set forth regarding systematic reviews (Higgins, 2011). Primarily, having a pre-determined protocol, method, inclusion and exclusion criteria is vital.

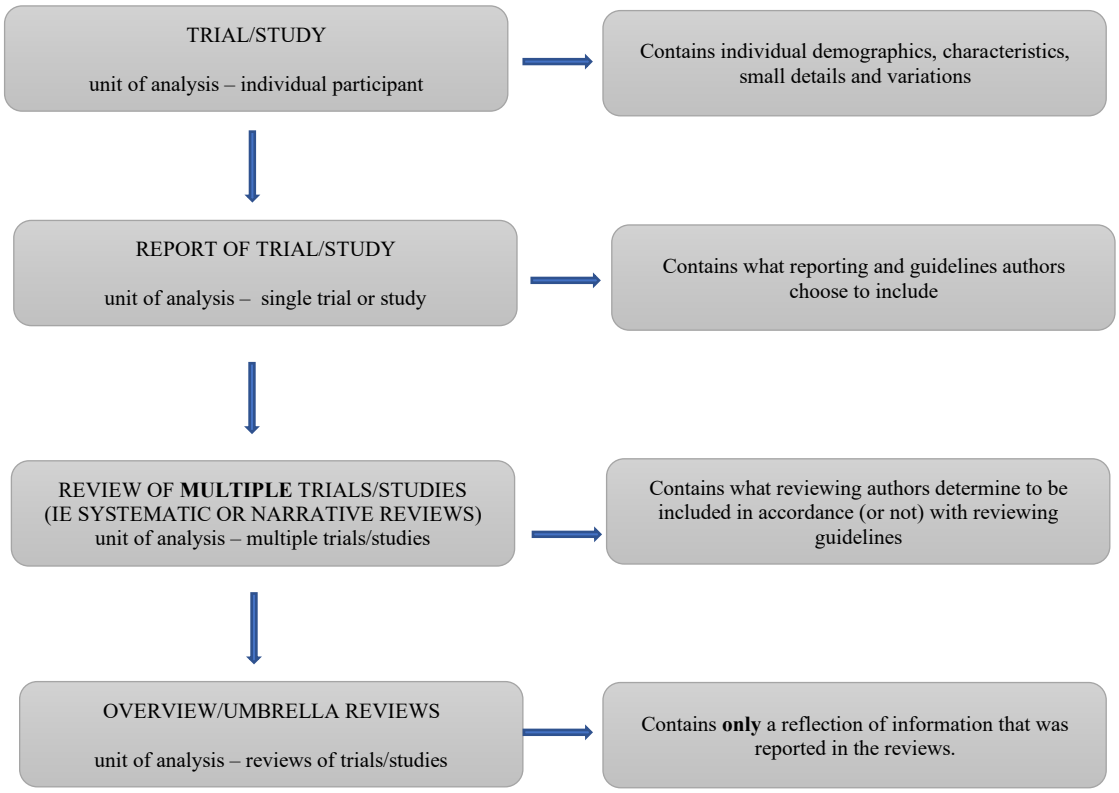

Figure 2. Overview flowchart

While there is a very well-established and consistent rationale for the completion systematic reviews and meta-analyses, the purpose, scope, and usefulness of an overview is less known and agreed upon. Systematic reviews have long been the standard for ascertaining whether information is consistent and generalisable across different settings, populations, and treatment modalities. Practitioners, stakeholders, providers and policy makers have large amounts of data and research available to them, it is unrealistic to expect that the time and resources necessary to peruse, understand, and evaluate this data will be set aside, and having reliable information regarding the evidentiary status of an intervention available in one comprehensive document is a useful and valuable element of an overview.

2.4.2 Assessing Bias in Overviews

It is critical to decision making and treatment planning to have the most unbiased and complete evidence possible. Methods employed by reviewing authors are designed with this goal specifically in mind. The practice of double-coding is one such tool, as is the increasingly common team approach to research and systematic reviewing, which will engage an often interdisciplinary method that draws upon the different expertise of different contributing authors. Uttley and Montgomery (2017) point out that including the perspectives of users and non-conflicted stakeholders may also be beneficial when conducting reviews; necessarily, the authors draw attention to the importance of always paying close attention to *who* is conducting the review, and *why*, and while a team approach is advocated, the members of that team must be chosen and vetted carefully and with the highest level of transparency possible. Allegiance bias, specifically, may be more easily avoided when more than one author engages with the topic, as differing perspectives and multiple attitudes towards the research question will be represented. However, it should be noted that, according to Eisner (2009), if even *one* member of a team has a vested interest in the outcome or conclusions of the work, it becomes vulnerable to allegiance bias. It is through critical analysis and synthesis of studies that the extant research in any given area - which on its own may be unmanageable or difficult to draw conclusions from - can be presented in a single cohesive and reliable document. Meta-analyses go one step further, possibly increasing the statistical power and effects of various interventions. The methodology and guidelines used in both systematic reviews and meta-analysis are designed specifically to reduce the incidents of the five main forms of bias described in detail in sections 2.2.1 to 2.2.5 and to be as reliable as possible in their conclusions.

Much of the literature surrounding the execution of overviews suggests that overview methodology is in many ways similar to that of a systematic review, and that these methods will be beneficial toward lessening the occurrence of bias (Pieper et al., 2012). When assessing the included reviews, the distinction between *methodological* quality and *reporting* quality is paramount, the first being related to how well the review was conducted, and the second refers to how well the findings are reported within the review; there is considerable debate on what means are appropriate to measure and ascertain methodological quality. Also vital to a complete an unbiased overview is the inclusion of reviews that may be seen as of poor quality (Caird et al., 2015). To leave out these perhaps lower-quality reviews would in effect be a form of selective reporting and would give an incomplete and biased picture of the evidence.

A number of frameworks have been developed in an effort to consistently and adequately measure and assess methodological quality. AMSTAR (Assessing the Methodological Quality of Systematic Reviews) was developed by Shea and her colleagues to build upon and refine the pre-existing tools for this purpose. Shea et al (2007) combined the enhanced Overview Quality Assessment Questionnaire (OQAQ) containing 10 items and a checklist created by Sacks et al (1987) containing 24 items with an additional 3 items based on advances made in the field and including language restriction, publication bias and publication status (Shea et al., 2007). The resulting instrument contains 16 questions relating to research design, literature review quality, included study descriptions, protocol information, risk of bias, funding sources and synthesis methods. AMSTAR is considered to be a useful tool, but in order to adequately complete the checklist, included reviews must report on a number of elements that may not be, for any number of reasons, consistently reported. Often AMSTAR cannot be used because of inadequate or incomplete reporting by review authors (Shea et al., 2007).

The Critical Appraisal Skills Program (CASP) has developed a series of tools designed to appraise different types of research. The ten-item checklist specific to the understanding and appraising of systematic reviews, broadly covers the following three areas: 1) are the results of included studies valid, 2) what are the results? And 3) will the results help locally? CASP is not a scoring system so much as a guide for thinking about reviews in a structured manner, or in a more workshop-oriented setting (CASP, 2018).

2.4.3 Strengths of Overviews Compared with Systematic Reviews

Systematic reviews do have certain limitations, one of which may be that they tend towards being restrictive in focus and do not capture evidence concerning a specific intervention's different applications (Hartling, et al., 2014). A systematic review is by design intended to answer a specific (and often narrow) question. They have a restrictive and finite nature, which may not always be the best possible means for approaching a broader research question. A Cochrane-compliant systematic review will normally contain only RCTs, and it will be conducted based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Higgins, 2011). A Cochrane Systematic Review is one of the four sub-types of reviews accepted and supported by Cochrane, and attempts to gather all the empirical evidence that fits with a predetermined eligibility criterion in order to answer a single research question. The Cochrane handbook is anchored by Methodological Expectations of Cochrane Intervention Reviews (MECIR), to which all protocols, reviews and updates comply. The handbook lays out in detail each section of a review, including what that section may or may not contain, how to construct the appropriate tables and figures, and all necessary headings and subheadings; authors can find pertinent, directive and clear guidance relating to each step of the reviewing process. The mandates put forth by Cochrane are intended for both authors and users and are intended to provide clarity and certainty of what to expect in reviews (Higgins, 2011).

There is no room for deviation, and perhaps this can be viewed as inflexible or strict. Cochrane reviews are designed specifically to promote evidence that is of high quality, free from conflicts of interest, accessible, and unbiased. There is an exhaustive and detailed manual for completion of a review, and functions as a sort of 'cookbook' for reviewers; a Cochrane Systematic Review uses "explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made" (Antman, 1992; Higgins & Green, 2011; Oxman, 1993).

Not all systematic reviews are published with Cochrane, it may be that authors use Cochrane principles and guidance, but chose to broaden inclusion criteria or eligibility; in certain cases, the research question may be better answered through a narrative review which, though often thought of as theoretical or contextual, can actually be a broader and more expansive take on the systematic review, being more flexible, containing different and broader inclusion criteria that will capture trials and studies that are not of only randomized design or were conducted prior to the advent of reporting guidelines. This approach may, in many cases, give a more complete and unbiased representation of the evidence. Some questions are not best answered through a systematic review, and another possibility for researchers who are attempting to surpass the limitations of systematic reviews is to undertake an overview of reviews. Three separate functions have been put forth for overviews. These three possible purposes as outlined in Ballard and Montgomery (2017) are as follows:

(1) the identification of gaps in literature wherein a number of studies exist but synthesis has not been performed,

(2) comparison of existing reviews, and

(3) to provide a summary of the current evidence within current systematic reviews, which may or may not include synthesis of this evidence.

It is dependent on what the intended goal of the overview is, a question which should be posed at the outset of the undertaking, which determines what approach to take to the process. The three possible reasons are not to be considered together, as each presents a different question, approach and methodology (Ballard & Montgomery, 2017; Pieper et al., 2012).

It has been suggested that overviews may be helpful when different reviews come to different conclusions, or identifying if this is in fact the case. This is confounding, and is due to any number of factors, including differing review methodology, inclusion criteria, possible issues with bias and inconsistent standards of measurement or identification of outcomes. It may also be that studies contain vastly different populations, or differences in outcomes may be attributable to other contextual factors such as location, socio-economic status, race of participants or the training of providers. Without further examination, there is no reliable means of understanding or justifying the evidence base of the intervention. The problem emerging from these issues requires the kind of clarification that an overview is in many cases able to provide. Through the rigorous evaluation and assessment of existing reviews important issues around the research and reporting of the intervention came to light. Resolving discrepancies between reviews is not the aim or function of the overview. Rather, the overview served to present the radically disparate and varied information in a comprehensive and understandable manner. While this method may not be as nuanced as the reporting of a single trial, or even a systematic review, it does serve as a "friendly front end" which collects and presents information relating to an intervention, and informs decisions relating to this intervention (Ballard & Montgomery, 2017; Cochrane Comparing Multiple Interventions Methods Group, 2013).

It has been argued that an overview, when conducted rigorously, may in fact be more able to identify the possible biases, reporting standards and theoretical implications than a systematic review (Hartling et al., 2014). Overviews present one potentially manageable and accurate means for decision makers, stakeholders and practitioners to understand and evaluate their decision making relating to a specific intervention.

2.4.4 Limitations of Overviews

It must be kept in mind that overviews are limited by the quality of those reviews contained within them, which are in turn informed by the studies contained within *them*. Overview reviewers must always be mindful of these limitations and present the results accordingly: with as much transparency and evaluation as possible. In order to achieve this, there must be an a priori protocol that is available and when it is deviated from, these changes must be made publicly. The data and reports must be easily accessible and published in an open forum. Additionally, where there are questions regarding data extraction, coding, search strategies or other details, this information must be readily and easily provided to querents. These systems are all put in place, once again, to reduce incidence of bias as much as possible.

Systematic reviews, meta-analyses and overviews contain little about the implementation elements of an intervention. Consideration of the dosage, delivery, uptake and context is often critical in determining whether an intervention works, and for whom it will be advantageous or positive. These elements, as well as a number of additional mediators and moderators, vary considerably across studies, and often have a significant impact on the outcomes of an intervention. In order to adequately understand the mechanisms and processes of an intervention a completion of an implementation analysis becomes necessary.

2.5 Implementation Evaluation in General

2.5.1 Purpose of Implementation Evaluation

Critics of Evidence Based Practice in particular have stated that it simply does not work in the so-called 'real world' (Macintyre and Petticrew, 2000). This could also be deemed the 'ivory tower' argument (Sackett et al., 1996; Gibbs and Gambrill, 2002). In short, the academics or scientists performing research and delivering interventions are far removed from complexity, difficulties and inherent chaos of the real world. So it may be that an intervention, when delivered by its designers, in as close to ideal circumstances as possible, with enough resources and support, can demonstrate large effect sizes and excellent results – results that could be seen as unrealistic. Determining whether an intervention 'works' or not is not the only question we should be asking. It is equally crucial to know for whom it is effective, and under what circumstances, contexts and conditions. Implementation science is one of the areas that delves into these issues, and is broadly concerned with why, how, and importantly *if* interventions are delivered consistently and as intended. The field contains within it the investigative methodologies that contribute to understanding why differences exist between primary trials as well as elements and frameworks that will contribute to, and promote, the uptake of quality research and evidence into common practice (Ghate et al., 2018; Mihalic, 2004). Implementation science and evaluation seeks to assist in the translation and adoption of research into practice and a more in-depth and detailed understanding of what influences the outcomes of an intervention (Elliot & Mihalic, 2004). The field is constantly being developed and expanded in order to better understand what interventions work for whom, and how to better deliver the best practices possible for a given population or problem.

2.5.2 Fidelity

It is acknowledged that understanding and remaining committed to the intent and concepts of an original intervention is critical to its success (Bonell et al., 2012). Also vital is that

researchers fully understand and maintain the processes that make an intervention possible and successful across different settings. In order for this to work, interventions should be grounded in a valid theoretical concept and mechanisms related directly to the theory of change as well as a consistent and high-quality evidence base. It is sometimes the case that fidelity can be measured alongside the evaluation process, through frameworks that also heed factors that are believed to influence implementation and outcomes. This process allows for implementation elements to be adequately tested, thus changes can be made when necessary and delivery can be improved. Some program designers attempt to ensure fidelity through the strict manualisation of the intervention. Interventions such as Multi-Systemic Therapy (MST), Incredible Years and FFT are examples of this approach to fidelity, and the manuals of necessity contain within them the activities, techniques and methods specific to the intervention, and serve as a guide and roadmap to their delivery (Alexander et al., 2013; Henggeler et al., 2009; Webster-Stratton, 2006). The purpose of this approach is to remain consistent with the original design of the intervention and to enable fidelity even when delivered by people other than its originators or in a different context.

The manner in which an intervention is delivered, that is, the fidelity to the core components and activities integral to the design of the intervention, is often seen as being at the core of whether or not the delivery is successful. Fidelity, in this context, refers to the manualized, structural and specific activities designed by an intervention's developers. This approach to implementation fidelity is a focus on the *form*, and while this can be measured through the use of process evaluations, there has been an emerging concept around whether or not this is in fact the best and most useful means of understanding fidelity. Recently there has been an emergent debate around whether the focus of implementation fidelity should instead be

on the *function* of an intervention, which may be defined as the "mechanism of change within the programs' theory of change" (Van Urk, et al., unpublished manuscript). It may be that through this different lens, program fidelity can be better understood, and programs delivered in a manner that is both true and consistent with the underlying change mechanisms, and also appropriate and possible in different settings or with different populations than the original context (Hawe et al., 2004). It must be kept in mind that interventions are very rarely delivered in strict, standardized conditions when they are taken up in real world settings. In order to maintain program fidelity, it may be useful to understand the foundational concepts and causal pathways underpinning the intervention and focus on this more functional approach to delivery.

2.5.3 The Oxford Implementation Index

Currently the guidelines in place such as the Cochrane Handbook and PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) statement do not contain within them enough information relating to the best method of gathering data relating to implementation, which we know is often at the heart of variable outcomes across trials. In an effort to remedy this situation, the Oxford Implementation Index was developed as a framework to be adopted and modified as needed; the Index contains four distinct domains, which are broadly defined as intervention *dose, delivery, uptake* and *context* (Montgomery et al., 2013). Within the FFT analysis, *dose* referred to how many sessions were given, and the cumulative amount of treatment time; *delivery* referred to the engagement process, and the beginning stages of FFT treatment; *uptake* presented difficulties with measurement, as it would normally refer to how much of the intervention participants would in fact be present for and engage with, or the difference between sessions provided versus sessions completed, which was not measured in any study. Finally, *context* referred to where the study was located, the characteristics and environment it was conducted in such as urban versus rural, in home versus in office, and the

demographics of study participants. Each of these domains may contain within it any number of different facets. Reviewers are urged to determine which characteristics are most germane to their topic and to modify the implementation index accordingly.

While the Index does provide a useful framework, it must be kept in mind that it is just that – a framework. It may assist in achieving a sound and systematic means of evaluating implementation data, but it is by no means exhaustive. Currently information relating to implementation data is not consistently reported in studies, and when the data are not present, the Index becomes less useful and there may be gaps in the information (Montgomery et al., 2013). It is necessary for trialists to continue to incorporate more complete and relevant implementation data which in turn will allow for a more complex understanding of intervention delivery and implementation. Reviewers are by necessity limited by what information is available to them, and an increased awareness of which elements are important for reporting will elevate the quality and reliability of systematic reviews.

2.6 This work: Overview of FFT

2.6.1 Rationale

There are over four decades of research into the method and use of FFT. And while the research and evidence surrounding FFT is in fact quite disparate, it is presented by the purveyors of the intervention as almost entirely positive and effective (Alexander, 2013). The Office of Juvenile Justice has published policy briefings and guidelines that refer to a strong and complete evidence base that demonstrates the efficacy and appropriateness of FFT for the treatment of behaviour disordered and delinquent youth, but their assessment of the evidence is incomplete and potentially biased, in part because they are limited to the published reviews and trials of FFT, many of which contain flaws or allegiance biases and do not specifically consider the potential for harm.

The impetus to use overview methodology in approaching the evidence around FFT stemmed in part from the disparity between the results found in the available reviews. The current evidence did not present a consistent picture, and an overview was conducted in an effort to better understand these disparities.

A closer examination of the literature suggested that FFT is not as effective as some reviews may indicate, and it became evident that a majority of the reviews, and studies contained within them, have been authored or conducted by those who may have a considerable bias (Alexander, 2013; Alexander & Sexton, 2002; Waldron & Turner, 2008). Overviews are an excellent tool for reduction of this form of bias, so it was decided to conduct an overview of the evidence in FFT.

When approaching the question of FFT's effectiveness, it became clear that the current research base would not be best developed and built upon by another systematic review or meta-analysis, in part because of the limited number of RCTs examining FFT. In the hierarchy, an overview is placed next to, or on the same par, as a meta-analysis, which is a more recent addition to the hierarchy and is placed above systematic reviews. In the Cochrane Guidelines, a strict overview can contain only Cochrane systematic reviews, but in the case of FFT this inclusion criteria would be far too restrictive and limiting and would fail to capture a great deal of vital information and evidence. An overview of reviews that includes narrative reviews would be a far more appropriate, useful, generalisable and applicable means of understanding the evidence. An inclusive overview design is able to capture information that would be of necessity left out from different methods and designs. A systematic review would not contain a number of studies because of their failure to conform with current inclusion criteria laid out by the Cochrane Collaboration, which includes that only studies with an experimental design may be

part of a systematic review; because of this a meta-analysis would *also* fail to include a number of reviews and reports. Even when examining reviews, each review may arrive at an entirely different conclusion, for any number of reasons, including how the authors determine inclusion and exclusion criteria, research design differences or author biases. Thus, in order to give the most complete and inclusive picture of the evidence, it became clear that overview methodology must be utilised.

Rather than narrow the inclusion criteria to only reviews that complied with the Cochrane Handbook guidelines, the overview examined narrative reviews and reviews authored before the advent of Cochrane (or other) systematic reviewing guidelines. And while the Cochrane Collaboration's guidance for the completion of overviews posits that only Cochrane systematic reviews should be included in overviews, this approach may leave out any number of quality reviews and present an incomplete picture of the evidence. Despite research suggesting that Cochrane reviews are of higher quality than others, the broader approach taken here aims to increase the validity, scope and value of an overview (Ballard & Montgomery, 2017; Pieper et al., 2012).

2.6.2 Inclusion criteria: Population, Intervention and Outcomes

The population of interest in the overview is youth aged 10-18 years with behavioural problems; the intervention is FFT; the comparator is alternative or no intervention; and objective or subjective outcomes were included.

Efficacy and effectiveness are distinct issues; the efficacy of an intervention refers to whether the intervention produces the expected or hoped-for result under ideal or 'laboratory' conditions. Effectiveness trials, on the other hand, are designed to measure whether the intervention will produce positive results in real world settings. The efficacy of any given

intervention should have already been established prior to undertaking trials meant to determine its generalizability and translation into wider practice (Higgins, 2011).

When determining FFT's effectiveness, it is necessary to examine whether the treatment had a positive or desired outcome on any of the three core outcomes as defined by the FFT literature. These three outcomes are:

1) reduction in recidivism by the juvenile;
2) reduction in substance use and misuse and;
3) reduction in out-of-home placements in punitive, treatment or protective settings.

Even something which at first blush may appear to be objective, such as recidivism, is in fact rather complex. Recidivism is basically a catch-all term used to refer to someone who, after receiving a criminal sanction, goes on to commit another infraction. The UK Ministry of Justice uses different methods for counting and reporting recidivism rates in adult and minor populations, as does the US, but both draw attention to the fact that true rates are likely much higher than what are reported, as the means for detecting recidivism are inaccurate. Not all crimes are recorded in a central database, and not all records are collated in a consistent manner (www.gov.uk; www.nij.gov). In the United States, there are 5 points which are typically used in the measurement of recidivism, these are: arrest, filing of charges, adjudication or conviction, commitment to a juvenile facility, and commitment to an adult facility. To further confound the matter, where the records are drawn from has an impact on the measurement. Court records often do not include any record of the arrest itself depending on whether charges are filed. Also, a status offense (any offense which pertains only to minors such as truancy or running away from home) may not result in sanctions, or it may not even make it into the final court database. Often the more liberal courts prefer to be informal when dealing with minors, so as to keep a young

person's record as clean as possible and in an effort to rehabilitate rather than punish; in this case, again, there is no way of accurately assessing recidivism (CJCA, 2018).

What may not be included are detentions not resulting in arrest, warnings, probation violations, misdemeanours or self-reported criminal behaviours. Additionally, after a certain amount of time, a criminal act will no longer be considered recidivism, but a new action. What adds even more to the confusion is that, in the US, each state has a different system in place, a system that, for example, means that in Texas a 16-year-old's case will be handled by the criminal court, but the same case would be in family court if he or she resided in Virginia. Consistency of measurement is nearly impossible, there exists no true standard or definition for what recidivism is, much less how to collect data relating to it or how to measure it. Illustrative of this problem is that, using the average of recidivism rates for a number of states, the calculated national rate could be anywhere from 25% to 55% (Snyder & Sickmund, 2009).

In the case of FFT research, reports of recidivism must be based on court documents or official records, as opposed to self-reported criminal acts, but the definitions of recidivism vary between studies, and often the definition of what was counted as recidivism is not included at all. When a study such as Gordon (1995) compares rates of FFT completers to that of the state average, how valuable and accurate is this information, really? Especially if there is no transparency around the issue. Substance abuse, also, may be measured as failed urine tests, or subjectively by self-reported or parent reported use. Out of home placements, despite their being labelled by FFT literature as an objective measure, may also be reported only when it is deemed 'official', or state mandated, which does not capture runaways or youth who choose, without state interference, to live with family or friends. In this way, even the core outcomes of FFT are not consistent across studies. While official records do offer a good starting point to

understanding efficacy, they are not always sufficient when trying to understand how or if the intervention is successful, and more importantly perhaps, if the recipients find it acceptable and useful for them and their families. There is a vast difference between success 'on paper' (i.e. no failed drug screens, no arrests, no suspensions at school) and success in a truer and more complex manner. While it is clearly important to gather information relating to what are commonly – perhaps inappropriately – known as objective measures, it must be kept in mind that even these measures are not entirely objective or consistent, and they may not be an honest or accurate reflection of what is actually occurring.

In terms of the harms of interventions, commonly researchers present the possibility of harm narrowly, as either direct psychological or physical injury resulting from an intervention. In the case of complex psychosocial or behavioural interventions, this is often expanded to include adverse effects resulting from the treatment, such as, in the case of Cambridge-Somerville and Scared Straight, increased likelihood to engage in criminal activity (McCord, 1978; Petrosino et al., 2013). It is necessary, when assessing the possibility for harms to consider this more broadly and to include such harms as opportunity costs, which occur when spending goes towards the support and uptake of an intervention that has uncertain or minimal outcomes rather than alternative and perhaps proven programmes (Bonell et. al., 2012).

In the case of an intervention such as FFT, which is intended specifically to achieve a reduction in criminal behaviour, substance abuse and out-of-home placement, yet another layer is added. Any intervention meant to achieve a reduction in criminality or behavioural misconduct has the potential to harm not just the individual or family, but must be framed in terms of potential harms to those who may be victims of criminal actions; harm must in this case also include the future social and economic costs of crime and substance abuse (Petrosino, et al., 2013). In this

manner if a harm is caused by a criminological intervention, the harm extends far beyond the individual.

2.6.3 Assessment of Bias in the Overview

For the purposes of the FFT overview, it was determined that AMSTAR and CASP guidelines would not, in themselves, be the most useful means of determining methodological quality of reporting within reviews in part because the variability in the included data within reviews presented problems with adequately completing the checklists. Thus, methodological quality had to be assessed on an individual basis, with each review being coded accordingly. Because overviews are, as noted, dependent upon the information in the reviews they contain, each review was assessed on a number of variables, including whether the reviewers had used any standardized or critical means of appraisal (and if so what that critical appraisal tool was), whether they had included the design of the included studies, whether they had any conflict of interest, reported or otherwise, what outcome measures were included and how these measures were determined, and whether there existed a protocol or a clearly defined exclusion and inclusion criteria.

It is impossible to eliminate all potential for bias, but it is important make every effort to do so. By utilising a set, open and transparent methodology, a clearly delineated search strategy and publishing an a priori protocol, most forms of bias can be lessened considerably. Additionally, ensuring that all available and relevant reviews are included in analyses is vital. The systematic and rigorous methodology utilised in the overview was designed in large part to decrease, as much as possible, biases and incompleteness.

2.7 This work: Implementation Evaluation

2.7.1 Rationale

Reviews of FFT have reported variable outcomes. For example, a review conducted by Gordon (1995) reports recidivism rates of 11% or below, while the highest recidivism rates (60% and above) are reported by Huey and Henggeler (2001). It must be considered that, even under the most ideal of circumstances, FFT is unpredictable. There is no way to deliver the model without being influenced by a myriad of factors; there is no way to prepare for or even guess at what might happen in a family home during a session, just as there is no way to control what happens at the youth's school, or in the community. Because of this, it is necessary to identify and understand the factors, no matter how apparently small, that may influence the outcomes of the intervention.

2.7.2 Methods

For the purposes of this analysis, the Oxford Implementation Index (Montgomery et al., 2013) was utilized as a framework with which to approach the central question of whether, and how, implementation of FFT across different settings may affect core outcomes. Data relating to the dose, delivery, uptake and context were identified and extracted from primary studies, along with information relating to supervision quality and amount. Through the coding of these data, implementation concerns and effects could be more apparent and identifiable.

2.8 Discussion

This chapter described the methods and rationale supporting the decision to eschew the more traditional designs of systematic reviewing, meta analyses, and process evaluation in favour of an inclusive overview design and an implementation analysis based on the Oxford Implementation Index. The nature and qualities of the available research relating to FFT did not allow for the aforementioned and more common design methodologies. An attempt to understand the research base of FFT in a truly unbiased and balanced manner informed the

decisions regarding research design. Bias in research is in many ways unavoidable and inevitable, especially when research is conducted by a single author, and can be separated into unconscious bias, which is often not recognized by the researcher, and conscious or intentional bias. Unconscious biases may be attributed to any number of factors, such as an individual's own training, education, experiences and context. People are often not impartial, despite their best intentions, and are far more likely to favour conclusions which they already believe to be true (Eisner, 2009).

Intentional misconduct or bias is in some ways more egregious, but it is also perhaps more easily identified, and may perhaps be less insidious a problem. Additionally, intentional misconduct is often thought to exist more frequently in drug trials or similar, as it often involves actual fraud or misrepresentation of findings. To assume this does not exist in behavioral or social intervention research is perhaps naïve. Bias – unintentional or otherwise, can occur at every level of research, with every decision, in this way, what begins as a small decision, a small bias, can lead to a large effect (Eisner, 2009).

In the case of this research relating to FFT, the methods were chosen specifically to offset the likelihood or possibility for bias in the reviewing or analysis of the intervention and implementation. The use of pre-determined protocols, rigorous scientific structure and the framework provided by the Oxford Implementation Index also aided in the reduction of bias. In addition to these structural and design decisions, it became important to present, with as much transparency as possible, the methods, means, and processes by which the research was undertaken.

The results of the two main research methods of this work are presented below. Chapter 3 reproduces a published article on the Overview of FFT, while Chapter 4 presents the implementation work, which has been submitted for publication and is currently In Press.

3 RESULTS OF THE OVERVIEW

The following pages reproduce a published article on the Overview of FFT:

Publication number: DOI: 10.1177/1049731518792309

Publication title: Functional Family Therapy (FFT)

Reference: Research on Social Work Practice (2019) Volume 29, 333-46.

Research Article

Functional Family Therapy (FFT) for Behavior Disordered Youth Aged 10–18: An Overview of Reviews

Research on Social Work Practice
2019, Vol. 29(3) 333-346

DOI: 10.1177/1049731518792309
journals.sagepub.com/home/rsw
SAGE

Clio Belle Weisman[1] and Paul Montgomery[1]

Abstract

Purpose: This overview of reviews analyses the existing evidence base of functional family therapy (FFT), which is a manualized, family-based intervention for youth with behavioral problems and their families. FFT has been implemented among youth aged 10–18 at risk of, or presenting with, behavioral problems such as delinquency, violence, substance abuse, sexual perpetration, and truancy. **Method:** A multipronged search was conducted across 15 databases, 10 websites, and expert contacts in February 2018. **Results:** The search yielded 159 hits of which 31 were included and critically appraised. Included reviews were published between 1986 and 2018 and the number of included studies ranged from 1 to 18 (including 20–5,344 participants). Main effects of the intervention on core outcomes (recidivism and substance abuse) were modest and out-of-home placement was not reported. Secondary outcomes were also modest but generally positive. **Conclusions:** Findings demonstrate that overall quality of reviews is low, which makes any certainties about FFT inconclusive; this overview provides a concise, valid, and methodologically sound synthesis of the research into FFT, which requires more rigorous investigation.

Keywords

delinquency, functional family therapy, youth, overview, umbrella review, FFT

Functional family therapy (FFT) is a manualized family-based intervention program for youth with behavioral problems and their families. Stemming from a systems approach to family therapy, FFT is a short-term intervention, consisting of approximately 30 hr of treatment. The intervention is designed to address family dysfunction by recognizing and modifying maladaptive family communication patterns, training family members to negotiate effectively, set clear rules and boundaries about privileges and responsibilities, and finally to generalize changes to community contexts and relationships. FFT has been used in the treatment of behavior-disordered youth and families for over four decades, but there are questions concerning the reliability and consistency of its evidence base. An overview of reviews or "umbrella review" gathers together the current research in one comprehensive and systematic document (Becker & Oxman, 2008; Higgins & Green, 2011). Unlike a systematic review, overviews are not necessarily limited in scope or focus, as they are not restrained to the primary study as the unit of analysis but are based on the systematic review; in this way, an overview is able to provide a broader type of information regarding a specific intervention addressing different outcomes and applications. An overview is an excellent means of addressing the following concerns: (1) use of the same intervention in the treatment of different problems or populations, (2) differing outcomes from the same intervention within the same issue or population, and finally (3) to determine whether there are any adverse effects resulting from the same intervention (Ballard & Montgomery, 2017). It is thus possible to shed light on the strengths and limitations of current research and reporting and allows for practitioners to easily and comprehensively make decisions based on the best evidence available. This overview addresses all three of the objectives regarding FFT and in particular to assess for risk of harm (Pieper, Antoine, Neugebauer, & Eikermann, 2014; Pieper, Beuchter, Jerinic, & Eikermann, 2012).

Description of Behavioral Problems in Youth Aged 10–18

Behavioral problems refer to a wide array of psychiatric disorders and psychosocial problems, both internalizing and externalizing, including oppositional defiant disorder, attention deficit hyperactivity disorder, and substance abuse disorders. About one third of children with childhood behavior problems develop conduct disorder (CD; American Psychiatric

[1] Department of Social Policy, University of Birmingham, Birmingham, United Kingdom

Corresponding Author:
Clio Belle Weisman, Department of Social Policy, University of Birmingham, Birmingham, B15 2TT, United Kingdom.
Email: cbweismanlcsw@gmail.com

Association, 2013). This group of disorders affects not only the youth and family but extends to have a negative impact on communities and society at large. In 2010, the Center for Disease Control and Prevention the United States reported 3.5% of children aged 3–17 were diagnosed with behavioral problems and 4.7% with illicit drug use disorder in the past year (Costello, Erkanli, & Angold, 2006). For the individual, these disorders are associated with an increase in both internalizing and externalizing behaviors and symptoms. Internalizing problems are associated with a lack of control or ability to recognize and cope with emotions, social isolation or withdrawal, and liability (McCulloch, Wiggins, Joshi, & Sachdev, 2000). Externalizing behaviors are likely to bring the attention of authority or parental figures, as they are overt in nature and may be hostile, aggressive, or destructive. Several studies have found that there is comorbidity between internalizing and externalizing behavior problems (Bornstein, Hahn, & Haynes, 2010; McWey, Cui, & Pazdera, 2010) which may lead to increased chance of self-harm or suicide. In addition to these symptoms, there is likely a higher rate of failure or underachievement in the academic arena and unemployment later in life (Petterutti, Walsh, & Valesquez, 2009).

Within the family setting, attention and care may be focused on the identified client to the detriment of siblings or the family as a whole. The balance within the family can be disturbed, possibly leading to divorce, abandonment, higher rates of alcohol and substance use, untreated mental health issues, and other maladaptive behaviors. Abuse, neglect, interpersonal violence, and sexual misconduct are possible results of family dysfunction, which is highlighted by the fact that in the United States, a person under the age of 18 commits 1 in the 12 domestic violence offenses that come to the attention of law enforcement (http://www.ojjdp.gov/ojstatbb; Prinz, 1988. Increased violence and delinquency contribute to growing unemployment rates and illegal activities such as substance use, underage alcohol consumption, and the sale of narcotics (Petterutti et al., 2009). These activities may lead to imprisonment that will in turn affect a youth's ability to contribute to the long-term economic productivity of the community and the nation as a whole. The United States spends approximately 5.7 billion dollars per year on the incarceration of minors (http://www.ojjdp.gov/ojstatbb).

Existing Responses to the Problem

A number of family-based interventions have been developed to prevent and/or treat behavioral problems among children and youth. Multisystemic therapy has gained favor in the court system as a mandated treatment for offending or high-risk youth and their families (Henggeler, Schoenwald, Borduin, Rowland, & Cunningham, 2009; Carr, 2000), as have the cognitive–behavioral approaches of aggression replacement training (Goldstein, Barry, & John, 1998), systemic family therapy, and moral reconation therapy (Little & Robinson, 1988); among these, FFT is one of the oldest and best known. Since its development in 1969, FFT has grown to over 270 active sites; while the majority of these are located within the United States, FFT has also been implemented in Belgium, the Netherlands, Sweden, Norway, The United Kingdom, Ireland, and New Zealand (Alexander, Waldron, Robbins, & Neeb, 2013). It was chosen for this overview because of its longevity and increasing uptake.

Description of the Intervention and Theory of Change

FFT reflects a core set of theoretical principles in which behavior is seen as a representation of the family relational system, that is, as indicative of the communication, patterns, and purposes of the family. Its developers describe the overarching goals of FFT as follows: *to change the maladaptive behaviors of youth and families, especially those who at the outset may not be motivated or may not believe they can change; to reduce the personal, societal, and economic consequences of disruptive behavior disorders and to do so with less cost, in terms of time and money, than many other treatments currently available* (Alexander, Pugh, Parsons, & Sexton, 2002). FFT uses a manualized approach and includes reframing, interrupting of negativity or blame, redirection of focus, interpretations of patterns of maladaptive behavior with links to emotions, a deepening understanding of actions, and communication training with focus on positive communication. It incorporates theories of information processing, social cognition, and the psychology of emotion.

FFT is a short-term (90-day), intensive, and comprehensive program designed for the treatment of behavioral misconduct in youth aged 10–18. It is a treatment modality rooted in family systems theory and cognitive–behavioral theory that can be delivered in clinical settings, school settings, or in the home. It is intended to address youth and families with a wide array of presenting problems including but not limited to criminal behavior, truancy, running away, sexual misconduct, substance abuse, risk of out-of-home placement, and as a reentry program for youth returning to the community following release from institutional settings. The program requires between 8 and 30 hr of direct service to youth and their families typically over an average of 12 home visits in 90 days. Over more than three decades, its designers have modified the original concept in efforts to increase effectiveness and uptake by service providers. According to the literature presented by FFT Inc., these changes have been made based on a "solid foundation of evidence" (Alexander et al., 2013).

Therapists are expected to adjust FFT to family members' capacities and to the specific problems they face. The focus of an FFT therapist should not be on one specific member of the family rather the therapists attempt to build an equal and balanced alliance with each member of the group. This technique serves to minimize blame and fosters an environment in which each family member feels understood and vital to the process.

Overview and Essential Implementation Elements of FFT

Core Components

The intervention consists of five major components and may include pretreatment activities: engagement in change, motivation to change, relational and interpersonal assessment, behavior change and generalization across behavioral domains, and multiple systems. These goals are accomplished in five ordered phases, the implementation of which is dependent upon the earlier phase; each has an assessment and intervention component directed at specific goals. The founders and advocates of FFT attribute its effectiveness to the careful sequencing of techniques, helped by the continuous assessment and intervention processes (Alexander et al., 2013). Details of the five implementation phases are as follows:

Phase 1: Engagement, wherein workers focus on forming a productive and balanced relationship with the entire family, which involves "maximizing factors which enhance the perception that positive change might occur (intervention credibility), and minimizing factors (e.g., poor program image, difficult location, insensitive referral) that might signify insensitivity and/or inappropriate resources" (Alexander et al., 2013, p. 11). This phase also requires a high level of engagement from the therapist including availability, contact with as many family members as possible, cultural sensitivity, and "matching" to the family unit. Therapists must maintain strength-based relational focus in this as well as all subsequent phases.

Phase 2: Motivation attempts to minimize hopelessness and to create a positive context. To develop or enhance family members motivation, workers identify and quickly begin to modify the pattern of changeable intrafamily risk factors, especially negativity, hopelessness, and blaming; "and initiate and/or strengthen intra-familial protective factors that can mitigate the effect of risk factors that cannot be changed" (Alexander et al., 2013, p. 12). Therapists focus on the relationships and interrupt blaming sequences while encouraging positive themes and reframing negative behaviors.

Phase 3: Relational Assessments are performed in order to understand and analyze information relating to the relational process of the family, as well as to plan for the following behavior change and generalization phases. Therapists continue to redirect the focus from individual problems to a relational perspective. This will aid in understanding negative or destructive behaviors in a new, more positive manner.

Phase 4: Behavior change is used to "develop long term behavior change patterns that are culturally appropriate, context sensitive, and individualized to the unique characteristics of each family member" (Alexander et al., 2013, p. 13). Workers focus on cognitive, interactive, and emotional issues; emphasize positive communication and parenting skills; and provide concrete resources that "guide and symbolize specific changes in behavior" (Alexander et al., 2013, p. 13). This phase aims to reduce intrafamilial risk factors and enhance protective factors. In this, the most directive phase of treatment, therapists may assign tasks, use technical aids, teaching, role-playing, training, conflict resolution skills, or other means for developing necessary skills.

Phase 5: Generalization is the final goal of treatment and focuses on the broadening of behavior change to other settings and social systems. This involves referrals, mobilizing community support systems, and modifying deteriorated family–system relationships with schools, probation officers, or other systems. "Generalization activities involve knowing the community, developing and maintaining contacts, initiating clinical linkages, creating relapse prevention plans and helping the family develop independence" (Alexander et al., 2013, p. 15).

After the conclusion of treatment, families are sometimes seen in "booster sessions"; these can be requested by the family or as a refresher if there is a relapse of maladaptive behaviors. A relapse may be signified by objective outcomes such as arrest or truancy or by family reporting a desire for assistance. These sessions are designed to reinforce the skills learned in the behavior change phase of treatment and to remind the family of their ability to achieve a higher level of functioning.

Underpinning all five phases is the concept of "obtainable change"; this refers to the setting of realistic and achievable goals. These goals are not merely those set by the therapist, authorities, probation officers, or other treatment providers but agreed upon and discussed within the scope of treatment. The youth and family are integral to the decision of which goals to pursue in treatment.

Assessment focuses on the functional nature of problems within the family rather than on a diagnosis and is described as a "continuous, multilevel, multidimensional, and multimethod process that includes individual, family, behavioral, and contextual factors" (Alexander et al., 2013, p. 22). The relational assessment examines not only the connectedness or relative autonomy of family members but also the hierarchical standing of each member in relation to the others. It focuses on why or how the family continues in its promotion and maintenance of problematic sequences, chains of behavior, events, and interactions; assessment will also identify key risk and protective factors for the family.

How the Intervention May Work

The model is suggested to be useful for complex and multidimensional problems because of its flexible structure and cultural sensitivity (Alexander et al., 2013). The attention to relational patterns and clear focus on family systems is thought

to be at the core of the intervention's success. As many FFT therapists provide services in the home, it is particularly suited to those clients who may be initially unwilling, unable, or otherwise unlikely to attend sessions in a clinical setting. Within FFT, there is a very strong stance against blaming and negativity. Neither individuals nor families are to be seen or treated as "bad" or wrong. Often the families treated by an FFT therapist may be described as resistant to treatment, but FFT attempts to overcome this through its use of the techniques discussed above. When there is a perceived failure or lack of improvement, it is the therapist who must identify ways to alter her approach and treatment of the family, not vice versa. FFT therapists attempt to provide a therapeutic experience unlike that to which most families are accustomed: An experience that focuses on alliance and fostering a cooperative working relationship. Families should feel engaged and aware rather than belittled or blamed for their problems.

Method

Objectives

To describe, assess, and evaluate the existing reviews of research on the effectiveness of FFT as used for the treatment of youth aged 10–18 presenting with behavioral problems. This overview is protocol driven, with a detailed research plan presented to a reviewing body at Oxford University Centre for evidence-based intervention, at which time a priori methods and outcomes were established and approved.

Search strategy for identification of reviews. Electronic searches were made of relevant bibliographic databases, government reports, and professional websites. Reference lists of articles were examined and experts contacted to search for the so-called gray literature. Searches for ongoing reviews or protocols were conducted including the Prospero database. There were no publication, geographic, or language restrictions. Searches covered the following sources up to February 5, 2018:

- *Biomedical sciences databases*: MEDLINE, EMBASE, PsycINFO, Cochrane Library; Prospero.
- *Social sciences and general references databases*: ASSIA (1969–), Dissertation Abstracts International (1969–), ERIC (1969), InfoTrac (1969–), ScienceDirect (1969–), Sociological Abstracts (1969–), Social Work Abstracts (1969–), Web of Knowledge/Web of Science (1969–), Social Sciences Citation Index (1969–).
- *Government policy sources*: U.S. Department of Health and Human Services, U.S. National Institutes of Health, U.S. Centers for Disease Control, Office of Juvenile Justice and Delinquency Prevention, U.S. Government Printing Office, U.K. Home Office, The Ministry of Health, Public Safety Canada, and Australian Institute of Criminology.
- *Personal contacts*: First author made contact with FFT developers, practitioners, and independent researchers to identify unpublished reviews.
- *Cross-referencing of bibliographies*: References in located reviews were checked to identify additional sources.
- *Search terms*: Search terms for MEDLINE (modified as necessary for other databases) were as follows:

    1. FFT (either as index treatment or as comparison arm) and
    2. (evaluation$ or outcome$ effect$ or review$ or analysis).

These terms were selected by considering previous reviews and via a library specialist at University of Oxford.

Criteria for considering reviews for this overview. Reviews that focused on randomized controlled trials (RCTs) and quasi-experimental designs which met the Cochrane Effective Practice and Organization of Care Review Group criteria were included in this overview. FFT must be reported, evaluated, and described in sufficient detail to enable analysis. Reviews had to clearly report on and evaluate at least one primary study of a specifically licensed FFT program or site. The reviews evaluated a specifically licensed FFT program or site and included the core elements needed for a program to qualify.

The review must have inclusion criteria for studies which compared a licensed FFT program to another service, alternative treatment, or no treatment. No treatment includes simple probation, which in the United States varies from state to state but is broadly defined by the Juvenile Law Center as any disposition involving the supervision of a delinquent youth in the community rather than in a secure confinement facility; commonly, it is comprised of approximately 85% checking in or monitoring and 15% education or mentoring services with possible elements of tracking devices, house arrest, or community service (https://jlc.org). Reviews may include any of the primary studies of FFT. Book sections, memos, government documents, and presentations were included. For complete screening forms, contact author.

Types of outcome measures reported. The three primary goals for both providers and recipients of FFT are reduction in recidivism rates, fewer removals from the home, and reduction of substance use or abuse (Alexander et al., 2013). Acceptable objective outcome measures include police and court records, out-of-home placement or the termination of parental rights, drug screening tests, and school reports of truancy, suspensions, or expulsions. Secondary outcomes are subjective and include parent and teacher reports of disruptive, acting out or unmanageable behaviors, self-reported delinquency, drug use, internalizing symptoms, peer relations, and family conflict levels. Family dysfunction, while a significant concern, is seen in FFT as residing at the core of more serious offending and high-risk behaviors and is not an outcome. FFT proposes in their training and dissemination materials that follow-up periods should last anywhere from immediately after the conclusion of treatment to 5 years postintervention, and the three primary

outcomes of FFT have been used as measurement of the intervention since its original development (Alexander et al., 2013).

Primary objective outcomes.

Recidivism defined as rearrest, incarceration, or reoffending. Recidivism reports relied upon court records, judicial determinations, probation reports, and conviction rates; self-reported or otherwise nonadjudicated criminal activity is not included in recidivism rates.

Substance use determined by standardized drug screening performed by agency, court, or probationary staff.

Placement out of the home or entrance into care in a group-home, foster care, juvenile detention, or secure treatment facility.

Secondary outcomes.

School attendance, as reported and documented in official school records.

Parental reports of externalizing behaviors of youth and family functioning as included in the integrated reported system contained in the FFT model.

Self-reported internalizing and externalizing delinquent behavior ideally collected through standardized instruments such as the Child Behavior Checklist, Behavioral Questionnaire for Children (BASC-2; Achenbach, 1991). Data drawn from self-reports authored by FFT Inc. included in the intervention manual were also eligible for inclusion (http://www.fftllc.com).

Reviews without at least one primary FFT study or without control comparisons were excluded, as were discussions of adaptations of FFT and programs similar but not explicitly stated to be an FFT program.

Assessment of methodological quality. The Assessment of Multiple Systematic Reviews (AMSTAR) checklist (Shea et al., 2007) for the evaluation of systematic reviews was selected, but there did not exist sufficient information contained within many reviews to adequately complete this assessment. Instead, reviews were assessed to determine the design of included studies, whether there existed any method of critical appraisal, and whether or not the authors had a connection to FFT Inc. that may result in a degree of bias.

Quality of evidence in included reviews. Quality of included reviews was assessed by whether or not included studies were randomized, length of follow-up, heterogeneity, imprecision, and whether trials were adequately powered. A number of the reviews include multiple trials that are not RCT, which places them at a lower level of quality. Additionally, evidence may be downgraded due to the fact follow-up periods are variable and often do not exceed 6 months posttreatment. Furthermore, a number of studies included in the reviewing literature contain sample sizes that are fewer than 100 participants, which leads to imprecision of evidence. The importance and impact of heterogeneity among participants within studies and between reviews is considered. Close attention was paid to whether there were enough common elements between reviews that allowed them to be discussed within this work. The issue of heterogeneity and review quality extends from the trials themselves to the type of included review, paper, or report.

Data extraction and management. A data collection form was piloted and all data coded and put into an extraction sheet. All data were independently double extracted and any disagreements resolved by a third party. For additional information regarding data collection, contact first author.

Results

Results of the Search

The search strategy, capturing titles, abstracts, and key words yielded a total of 159 references, considered as possible reviews or meta-analyses, with 31 duplicates and two registered protocols of reviews currently in progress. After screening abstracts, 31 were excluded as not being reviews of primary research and 91 identified for full-text retrieval. After examining the texts, 60 were discarded for ineligibility (identified as primary research, training manuals, not containing information relating specifically to FFT, or did not report results for FFT); 6 meta-analyses and 25 narrative reviews, totaling 31 reviews ($n = 31$), met all eligibility criteria and were evaluated for overall quality, scientific rigor, quality of evidence, and potential bias. See the PRISMA chart (Figure 1).

Included Reviews

There were 31 reviews or meta-analyses, chapters, or policy briefs which met inclusion criteria. All reviews met the qualifier of discussing only those programs that are licensed and clearly stated to be FFT. Programs that may contain similar elements to FFT but were not licensed as such were not included. Nor were reviews only discussing adaptations of FFT. All included reviews must cite and discuss at least one primary study of FFT in enough depth to justify inclusion in this overview. See Table 2 for list of excluded reviews.

Research has suggested that there is a distinct possibility that reviews conducted or written by investigators who have an allegiance to the program models they are investigating produce significantly more positive results than those conducted by investigators without such allegiance (Eisner, 2009; Shadish, Cook, & Campbell, 2002). The reported effect sizes of prevention and intervention trials have been shown to be noticeably larger when program developers are involved as opposed to those conducted by independent researchers (Eisner, 2009). Alternatively, it may be important to note that positive effects of an intervention during a study conducted by "developers—as evaluators" may be attributed to either (a) the possibility that the quality of implementation is of a higher standard when it is delivered by a developer of the intervention or (b) disparity in outcomes between studies delivered by

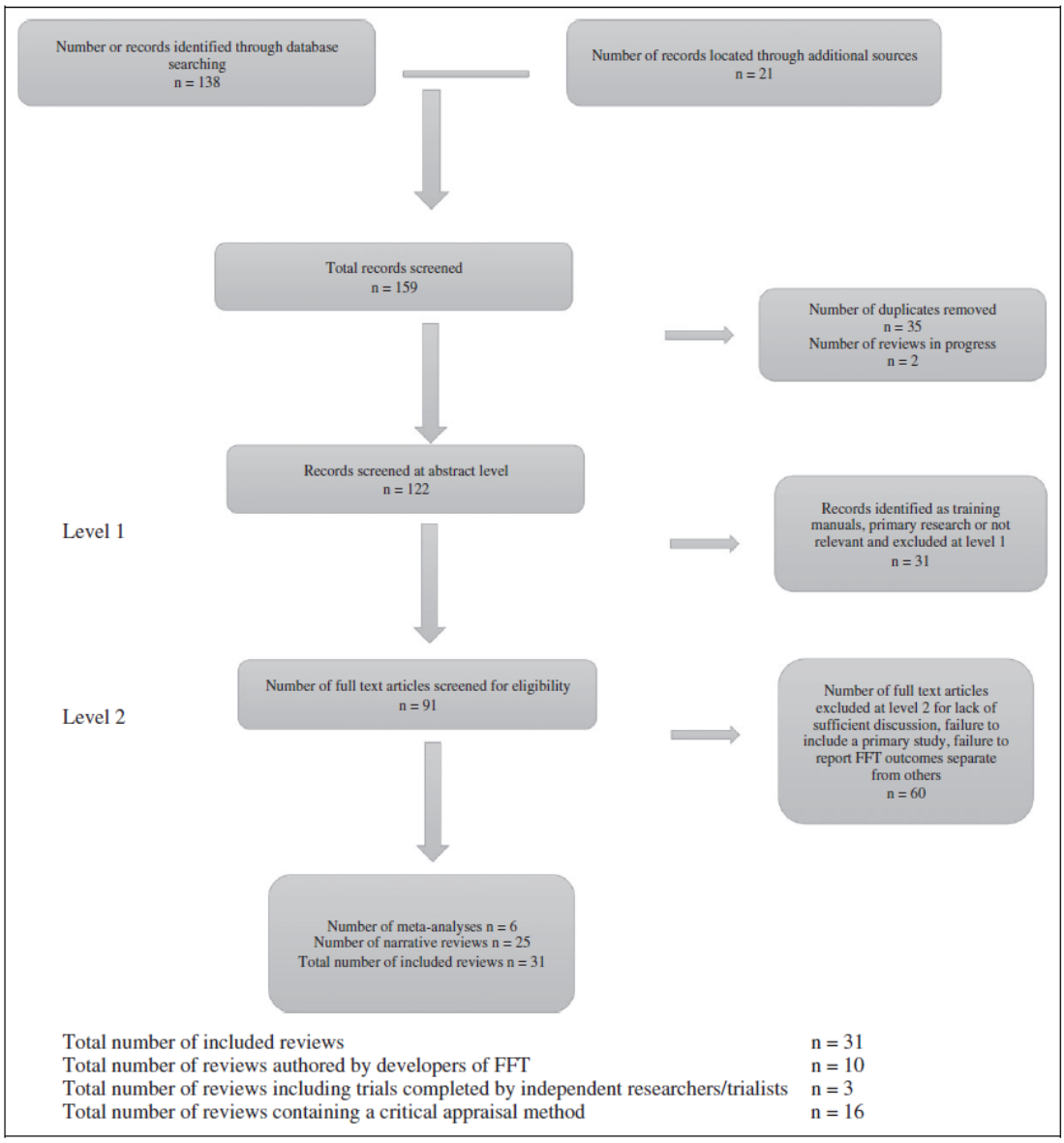

Figure 1. PRISMA flow diagram.

developers versus independent researchers that can be attributed to systemic bias relating to an ongoing conflict of interest by a champion of the program and possibly related to financial factors. The methods of this overview consider the presence of reviewer bias and study design, by comparing results of reviews published by FFT program developers and stakeholders to those obtained by others.

Design of included reviews. The trial design of included studies is not stated in the majority of reviews, and only 16 of the 31 utilize any standardized or peer-reviewed critical appraisal method (see Table 1).

Participants in included reviews. Only 10 reviews reference the age of trial participants and sample inclusion criteria, and only

Table 1. Design of Included Studies.

Review	Number and Research Design of Studies Included in Review	Total Number of FFT Participants	CoI—Involvement of Authors or Trialists With FFT Inc.	Critical Appraisal of Included Evidence (Grading of Recommendations Assessment, Development and Evaluation, CASP, Peer-Reviewed Tool)
Alexander (2000)	RCT-7	1,292	DEV	No
Alexander (2000)	RCT-6			
	Quasi-4	402	DEV	No
Alexander (2002)	RCT-6			
	Quasi-—4	1,959	DEV	No
Alexander (1998)	RCT-5			
	Quasi-5	1,399	DEV	No
Austin et al. (2005)	RCT—1	120	IND (Trials DEV)	Yes
Baldwin, Christian, Berkeljon, and Shadish (2012)	RCT-3	310	IND (Trials DEV)	Yes
Bender, Tripodi, Sarteschi, and Vaughn (2011)	RCT-1	114	IND (Trial DEV)	Yes
Breunlin et al. (1988)	RCT-3	143	IND (Trials DEV)	No
Diamond et al. (1996)	Quasi-4	241	IND (Trials DEV)	No
Dumas (1989)	RCT-2	107	IND (Trials DEV)	No
Farrington &Welsh (2003)	RCT-1			
	Quasi-1	138	IND (Trials DEV)	No
Filges, Anderson, & Jorgenson (2016)	RCT-2	225	IND (Trials DEV)	Yes
Gordon (1995)	Quasi-3	106	IND	No
Hartnett, Carr, Hamilton, and O'Reilly (2017)	RCT-11			
	Quasi-7	1,674	IND (Trials DEV)	Yes
Henggeler and Sheidow (2012)	RCT-4			
	Quasi-2	1,366	IND (Trials DEV)	Yes
Henggeler and Sheidow (2003)	RCT-1			
	Quasi-4	315	IND (Trials DEV)	No
Hogue and Liddle (2009)	RCT-4	359	IND (Trials DEV)	No
Huey and Henggeler (2001)	RCT-2			
	Quasi-1	189	IND (Trials DEV)	No
Mulvey et al. (1993)	RCT-1	40	IND (Trial DEV)	No
Robbins, Alexander, Turner, and Hollimon (2016)	OBS-12			
	Quasi-12			
	RCT-20	5,344	DEV	No
Sexton (2002)	RCT-5	181	DEV	Yes
Sexton (2004a)	RCT-1			
	Quasi-2	166	DEV	Yes
Sigrid, Qais, and Zepeda (2013)	RCT-1	135	IND (Trial DEV)	Yes
Stanton and Shadish (1997)	RCT-1	135	IND	Yes
Tanner-Smith, Wilson, and Lipsey (2013)	RCT-2	126	IND (Trials DEV)	Yes
Tarolla, Wagner, Rabinowitz, and Tubman, (2002)	Quasi-3	155	IND (Trials DEV)	No
Tolan (1986)	Quasi-4	267	IND (Trials DEV)	Yes
Vaughn and Howard (2004)	RCT-2	249	DEV	Yes
Waldron and Turner (2008)	RCT-3	341	DEV	Yes
Waldron, Brody, Robbins, and Alexander (2013)	RCT-6	555	DEV	Yes
Weisz, Hawley, and Jenson Doss (2004)	OBS-1	20	IND (Trial DEV)	No

Note. RCT = randomized controlled trial; FFT = functional family therapy; CoI = conflict of interest.

Table 2. Table of Excluded Studies.

Review	Reason for Exclusion
Axford (2013)	No primary study of FFT included
Aultman-Bettridge (2007)	Despite significant discussion no included study
Bachman (2010)	No primary study of FFT included
Carr (2009)	No primary study of FFT included
Farrington (2010)	No primary study of FFT included
Fonagy (2002)	No primary study of FFT included
Heekerens (1988)	Not reviewing FFT specifically
Hipwell (2006)	No primary study of FFT Included
Kazdin (2015)	Insufficient discussion of FFT, focus on psychological versus behavioral interventions
Lipsey (1998)	Does not report or discuss FFT separately from other family therapies
Stratton (2005)	No primary study of FFT included
Stratton (2011)	No primary study of FFT included
Underwood (2002)	No primary study of FFT included
Wiggins (2012)	No primary study of FFT included

Note. FFT = functional family therapy; CoI = conflict of interest .

5 contain information regarding race or gender of study participants. Contextual demographic information is often missing from the reviews.

In the meta-analyses, the authors briefly describe the inclusion criteria for the participants in the trials. Excepting Farrington (2003), the meta-analyses include youth aged 11–18, with a diagnosed substance use or abuse disorder. While age and sex of participants is included, there is no further information concerning core elements that may directly impact the outcomes of the intervention.

Primary Outcomes Reported

Recidivism as an outcome measure is reported in 16 of the reviews, which includes 11 studies containing a total of 1,427 unique participants. Ten reviews include a percentage rate for reoffending ranging from 11% to 67% recidivism in the FFT groups versus 36% to 93% in comparison and control groups, although definitions of recidivism vary widely across reviews. Sexton, Ridley, and Kleiner (2004), Sexton (2004b) and Dumas (1989) report "significant reductions" in recidivism and "less than half" the recidivism rate of the comparison or control group, without any specific data being provided. Diamond, Serrano, Dickey, and Sonis (1996), Huey and Henggeler (2001), and Chamberlain (1995) state that FFT results in reduced recidivism but provide no clear data nor specify the type and severity of crimes. Across the reviews, authors do not clarify which definition of recidivism is being used. Only four ($n = 2,629$) state what reports or records were accessed within the studies to gain this information: Alexander and Sexton (2002), Gordon (1995), Henggeler and Sheidow (2003), and Vaughn and Howard (2004).

Substance abuse is only reported in 10/31 reviews; six ($n = 1,846$) of these are meta-analyses focusing on cannabis or substance abuse. The included meta-analyses are the only reviews that report effect sizes within them. The effect sizes reported range from a Hedge's g of -1.18 to 0.21, with three meta-analyses reporting nonsignificant effect sizes (Baldwin, Christian, Berkeljon, & Shadish, 2012; Bender, Tripodi, Sarteschi, & Vaughn, 2011; Vaughn & Howard, 2004). Of the remaining four, Hogue and Liddle (2009; $n = 349$) state that there was "an improvement in drug use" (p. 18); Vaughn and Howard (2004; $n = 249$) report a 50% reduction in the drug severity index; and Waldron and Turner (2008; $n = 341$) state that the treatment group "showed significant reductions in substance use at post and/or follow-up" (p. 248).

Out-of-home placements are not reported as an outcome of any study despite claims that FFT is an effective tool in the reduction of out-of-home placement for children and youth who are at risk of removal (Alexander et al., 2013).

Secondary Outcomes Reported

School attendance is not discussed or reported in any review of FFT, nor is there any reported collection of data concerning school attendance. This information is included neither in the included reviews nor in the registered protocols of FFT reviews still in progress. Thus, no judgment can be reached regarding whether FFT will result in better (or worse) school attendance and truancy.

Parent reports of family functions, interactions, and youth-related conduct issues were not routinely included, with no mention of clinical services system reports, no questionnaires, and no other data collection tool utilized in any of the included reviews. It is not possible to determine whether it is the trials themselves or the author reporting them that did not collect this information. The FFT model has embedded within it evaluations and assessments relating to family functions, and it is striking that these data are not included in the reviews of the research.

Self-reported internalizing and externalizing delinquent behaviors are also absent from reviews. Again, FFT has a method of collecting these data, but it was either not collected or not reported. Additionally, reviews did not contain the use of, or reference to, other measurement scales such as the BASC-2 or other peer-reviewed depression or behavioral scales. Information relating to self-reported criminal activity, depressive symptoms, family conflict, school-related behaviors, and other behavioral- or conduct-related issues are not present.

Other Outcomes Reported

Earlier reviews of FFT, which reference the initial studies such as Alexander and Parsons (1973; $n = 20$); Klein, Alexander, and Parsons (1977; $n = 86$); and Barton, Alexander, Waldron, Turner, and Warburton (1985; $n = 27$), report findings that are neither objective nor easily quantifiable. These outcomes are

described as "reductions in conflict level," a "longer duration of silence," and "decreased defensiveness." These outcomes, while appearing to be beneficial, are open to interpretation and do not appear to be based on valid, objective, or peer-reviewed measures.

Quality of Evidence in Included Reviews

Reviews, book chapters, and reports written by the creators of FFT or by individuals who are now or have been employed by FFT, Inc. may be at a high risk for bias. Only 5 of the 31 included reviews contain summary of findings tables with risk of bias assessments and only 16 include the use of a peer-reviewed critical appraisal method. These reviews generally report only the positive outcomes or measures while ignoring other less favorable outcomes. See Table 2 for complete listing of reviews and trial design of included studies.

Effect of Intervention

When examined as a whole, the reviews describe effects of FFT in inconsistent terms. Importantly, no review reports a directly harmful or detrimental outcome. A number of reviews, however, report no significant differences between groups in any of the outcome domains including family functioning or decreased internalizing and externalizing behaviors (Austin, Macgowan, & Wagner, 2005; Baldwin et al., 2012; Tarolla, Wagner, Rabinowitz, & Tubman, 2002; Tolan, 1986). Recidivism is measured by different means across studies, and the reviewing authors do not specify which measure or definition of recidivism they are using as an outcome measure. This makes any assumptions or inferences regarding recidivism extremely difficult to substantiate. The recidivism rates reported range from 11% to 67%, and Farrington (2003) draws attention to the fact that although low recidivism rates were reported, there were definite methodological problems with attrition and variable follow-up periods. The lowest recidivism rate was found in Gordon (1995), in which authors cite a quasi-experimental trial they had conducted in rural Ohio with court-mandated youth ($n = 27$). The highest rates of recidivism (60% or above) are reported by Huey and Henggeler (2001; $n = 189$); Mulvey, Arthur, and Reppucci (1993; $n = 40$); and Gordon (1995; $n = 106$) and refer to a trial conducted by Barton and colleagues in 1985 (Barton, Alexander, Waldron, Turner, & Warburton, 1985). Specifics of this trial are not included in the reviews.

Discussion

Summary of Main Results

This overview aimed to gather and assess the preexisting reviews of FFT and to present information regarding the model in a single comprehensive document. The narrative reviews of FFT did not suffice as a reliable source of information regarding the model and left a great deal of room for conjecture and uncertainty. The three primary outcome measures promoted by FFT Inc. have not been adequately addressed in the existing literature. This overview sheds light on this issue and highlights the importance of increasing involvement of those not directly linked to the development or dissemination of FFT Inc. Even with the evidence currently available, there is no indication that a course of FFT will not cause detrimental effects, or harm. Without this certainty, it seems unwise to advocate for the intervention without hesitation.

Recidivism is a complex outcome which is open to interpretation. It can refer to any number of outcomes including contact with the court system, rearrest, conviction, and self-reported criminality. A higher standard of proof is employed when obtaining a conviction versus an arrest. To measure recidivism, only as a subsequent conviction greatly skews the outcomes in favor of a lower recidivism percentage rate. Differing definitions of recidivism may lead to any reported rate, across the included reviews, becoming less meaningful as a result. The picture of evidence that emerges from studies employing completely different standards of measurement will be diverse; without identifying and defining these differences, we cannot reach valid conclusions. According to the National Institute of Justice and the Office of Juvenile Justice Prevention, a status offense such as truancy or probation violation can be counted as recidivism but should not be measured by the same metric as a felonious arrest. An arrest or conviction requires a far greater standard of suspicion and proof, whereas status offenses are most often reported by probation officers and do not contain antisocial or disruptive elements. These inconsistencies and lack of designation often leads to an overarching lack of clarity in the conclusions (https://www.nij.gov; Measuring Recidivism. (n.d.). Considering many reviews did not report this outcome, it becomes even more difficult to reach conclusions relating to this primary outcome, which is unfortunate since the intervention was in part specifically designed for this purpose.

Farrington (2003), a meta-analysis with a focus on recidivism rather than substance abuse, concludes that the two trials (Alexander & Parsons, 1973; Gordon, 1995) included do lead to positive results and a significant effect. Despite a better standard of reporting and analysis than many other reviews, the sample size in these two trials was 52 and 86, respectively (Farrington, 2003).

While FFT is considered to be a method of reducing out-of-home placements, no review reported on this outcome. It is perhaps possible that out-of-home placements increased in duration or frequency. Without clearly stating this result, it is unwarranted to assert that FFT reduces these occurrences as claimed in FFT site and training literature (Alexander et al., 2013; http://www.fftllc.com). For this outcome to go unreported and undocumented is troubling at least and possibly a significant failing of the current research and trials of FFT. Similarly, reduction in substance abuse is one of the primary outcomes as stated by FFT, and gathering objective data relating to this outcome would do a great deal in terms of strengthening the evidence around the claim that a course of FFT will lessen incidents of substance misuse.

School attendance as an outcome is also not sufficiently addressed in the literature, and it is striking that such an objective measure was not reported in the reviews. It is impossible to know whether these data were simply not collected, thus being a reporting error, or whether the reviewers chose not to include it in their results, because they judged it to be insignificant, they did not have access to it, or for any other reasons.

FFT does include in its design multiple methods for the reporting and rating of internalizing symptoms, but these records are not included by the reviewing authors. Additionally, there are additional standardized methods of measuring these symptoms that are either not employed by trialists or not reported on the results.

The current evidence suggests reviews may not include all available research and pertinent contextual information in analyses, resulting in an incomplete or skewed picture of the evidence. It is impossible, when examining the reviews of FFT, to establish whether the trials referenced are well designed, executed, and significant as the information necessary to make such judgments is too often absent from the review. Reviews are vulnerable to biases and "perspective" that may influence the reader to lean in one direction or another. Reviews do not examine any possible bias of referenced trials or include a risk of bias assessment, though one mentions that "most trials were performed by the developers of the model." Reviews must have a sound methodological basis for reaching any conclusions, an approach that is outlined and supported by both the Cochrane Grading of Recommendations Assessment, Development and Evaluation guidelines and AMSTAR (Higgins & Green, 2011; Shea et al., 2007).

Reporting guidelines that have been widely adopted were not employed. Consolidated Standards of Reporting Trials (CONSORT) has been developed, designed, and is being improved upon to assist in the consistent and standardized reporting of trials (Motgomery et al., 2018; Schulz, Altman, & Moher, 2010). Still, it would behoove researchers to adopt these guidelines due to the fact that, were trials more completely and rigorously reported, the standardized reviews of them would in turn be of a higher quality. If CONSORT guidelines had been utilized in the reporting of the FFT trials, current available research would be of greater value and substance.

Reviews of FFT often do not report or discuss any elements of implementation or contextual concerns, staff characteristics, or adherence within the trials. These issues are known to have a significant impact on the outcomes and efficacy of interventions (Montgomery et al., 2013a), and this information could be invaluable when determining the strengths and weaknesses of the model.

When describing sample inclusion criteria, 10 reviews (Alexander et al., 2002; Austin et al., 2005; Breunlin, Breunlin, Kearns, & Russell, 1988; Cottrell & Boston, 2002; Dumas, 1989; Gordon, 1995; Henggeler & Sheidow, 2003; Huey & Henggeler, 2001; Sexton, 2009; Waldron & Turner, 2008) refer to study participants, but the criteria are not elucidated well and include phrases such as "light to severe offenses," "habitually truant," "soft" juvenile offenders, and "serious juvenile offenders." There is no definition of these terms and no explanation of what level of behavior is present. There is a great deal of room for interpretation. Waldron and Turner (2008) report outcomes for youth diagnosed with "heavy alcohol abuse"; Dumas (1989) reports the participants were court referred; Austin, Macgowan, and Wagner (2005) discuss a sample that required a substance abuse diagnosis; and Breunlin, Breunlin, Kearns, and Russell (1988) report on a study with status offenders. These are definable categories which are far more objective and useful when examining the quality of design and methodology in the included reviews.

Contextual information such as socioeconomic status, number of family members, criminal backgrounds, and education has been shown to have an impact on effectiveness and outcomes of FFT. Thus, it is problematic that this information is not included by the review authors. Making well-informed decisions relating to the use of FFT for specific populations based on evidence is difficult if and when this information is not readily available or accessible. A high-quality systematic review of the evidence may help in remedying this situation, but for the purposes of this overview, such information was not readily available in the extant literature.

Outcomes and the Theory of Change

The 31 reviews of FFT included do not report consistently on the core outcome measures of FFT, and therefore, it becomes difficult to determine whether it is in fact an effective method of treating conduct and behavioral problems in youth. There is no clear linkage between the outcomes FFT Inc. promotes treatment for and the research outcomes. Of the six main outcomes described by the developers, substance abuse is reported in 7 of the 31 reviews and recidivism in only slightly over half of them ($n = 16$). This does not allow for any strong conclusions to be reached regarding the treatment's efficacy. But once more, it is not possible to ascertain whether this is due to possible design or implementation failure in the trials or to a lack of consistency and breadth in reporting.

Trials examining FFT have been completed almost exclusively by the developers of the model or by those who otherwise benefit from its use and uptake and have a significant stake in the success of the model. Even in instances of independent reviewing authors, the trials contained in these reviews were performed by developers of FFT. This may influence the outcomes in a number of ways. When a trialist is particularly familiar and involved with the intervention being researched, it may lead to increased fidelity and adherence to the original model design (Elliott & Mihalic, 2004). However, it has been indicated that studies conducted by investigators who have an allegiance to the program models they are investigating produce significantly more positive results than those conducted by investigators without such allegiance (Eisner, 2009; Shadish et al., 2002; Oxman & Guyatt, 1991).

There are many persuasive arguments relating to the need for standardized methods relating to the completion, evaluation, and reporting of trials, and these demands become

increasingly important and difficult when examining a complex intervention. Fidelity must also be examined not just as adherence to a manual or program guideline but fidelity to the underlying mechanisms themselves. Implementation concerns are not adequately addressed in the existing reviews; FFT is largely believed to be a very adaptable model, a belief that is not demonstrated in the research. It is unknown whether the trials themselves do not examine implementation or whether the reviewing authors did not gather this information. Both implementation and fidelity are major concerns when assessing the outcomes and overall effectiveness of a complex intervention. These issues cannot be adequately rated within the scope of this overview; the reviews included in this trial did not contain within them consistent or detailed information regarding duration of treatment, skill, and training of practitioners; context; and adherence to the model or attendance of either practitioners or recipients of the intervention. Without this information, it is not possible to ascertain what concerns and impact implementation may have on the outcomes. Such implementation data should have been addressed and reported in the original reviews; where there is no discussion of this, it was noted as a possible demonstration of failure on the part of either the reviewers or those running the trials to heed the importance of implementation fidelity (Montgomery et al., 2013b).

Overall Completeness and Quality of Evidence

The existing reviews do not offer sufficient evidence on which to base decisions concerning FFT. Of the 43 available randomized, quasi-experimental, and observational trials across 31 reviews which include a total of 4,607 unique participants fewer than half are identified, included, or discussed in the reviews. In many instances, the reviewers fail to specify populations served, participant characteristics, or outcomes measured in the trial. The reviews of FFT do not provide information such as whether or not a protocol existed or was registered, search strategy for identification of studies, methods used, or outcomes measured. This creates difficulties in the gathering of substantial information concerning the quality or completeness of many of the reviews.

Agreements and Disagreements With Other Studies or Reviews

Prior reviews of FFT present a picture that is incomplete, but largely positive. The reviews date back to the mid-1970s and contain varying degrees of detail, thoroughness, and complexity. Previously published reviews state with minimal hesitation that FFT is an effective and beneficial method of treatment for behavioral problems, CD, and substance abuse issues. It appears that this is a conclusion arrived at without the use of consistent or sound methodology. This overview employed rigorous methods of identification and evaluation of the existing reviews. Additionally, tables included in this overview were completed not through information provided in the reviews but by use of the original primary studies in an effort to gain a better picture of the evidence. The sound and comprehensive search strategy, methodological process of evaluation, and analysis methods arrive at a different conclusion; one that presents significant areas for further examination and research, as well as shedding light on some of the questions regarding efficacy of FFT across different populations and highlighting the need for implementation analysis.

Potential Biases in the Overview Process

It is impossible to eliminate all potential for bias, the authors have made every attempt to locate and evaluate all the relevant reviews, as well as double coding data and having a third party resolve any disagreements. There is also a risk of bias stemming from the absence of reviews still in progress; currently, there is a registered literature search and meta-analysis which is being funded by FFT Southwark with an unknown prospective date of completion. A systematic review of FFT protocol is currently being undertaken and is registered with the Cochrane Collaboration and when completed will be adherent to Cochrane guidelines.

This overview gathered and assessed the preexisting reviews of FFT and presented this information in a comprehensive and rigorous manner. It was determined that the existing reviews of FFT did not suffice as a reliable source for this information and left a great deal of room for conjecture and uncertainty. The primary outcome measures promoted by FFT Inc. have not been adequately addressed in the existing reviews.

Strengths and Limitations

Overviews aim to provide a comprehensive picture of the existing evidence for a particular intervention, by assessing and examining the existing reviews. This overview differs from existing reviews in that it meets the standards set out in the Cochrane handbook guidelines for rigor and quality. The methods used in searching, data extraction, and synthesis are clearly detailed, possible biases are reported, and there is a higher level of transparency than is present in the existing literature whether in reviews or overviews. Double coding was employed throughout the overview, and authors were contacted to confirm or clarify data when necessary.

Conclusion

This overview outlines the uncertain effects of FFT across different populations, and the existing reviews illustrate the fact that certainties regarding FFT are based on a tenuous research base. The initial research, effectiveness, and efficacy trials were held in Salt Lake City, UT. While there are some trials of FFT conducted by independent researchers ($n = 17$), the majority of research continues to be conducted by the developers and disseminators of FFT ($n = 27$). Additionally, many of the reviews currently available are authored by the

developers of the model. While this is to some degree expected, it is demonstrative of a need for independent researchers to conduct well-designed and well-reported trials and subsequent reviews of FFT.

Given these factors, it may not be advisable to continue adopting FFT without reexamining and testing the effects. It is difficult to determine the quality of any evidence regarding the efficacy and impact of FFT and it is necessary that this situation be remedied. Trials and replications should be performed and reported with greater stringency around guidelines and quality control. FFT has not yet been adequately proven effective across culturally disparate and varied populations. There is a need for well-powered trials that include all primary outcome measures of FFT this would measure out-of-home placements, and internalizing and externalizing behaviors; utilize a consistent measure of recidivism; and evaluate cost-effectiveness. It appears that in nearly 40 years of existence, there continue to be a number of unanswered questions regarding effectiveness and implementation. Continued research and investigation with a focus on implementation with a high standard of design will do a great deal to address these concerns.

Acknowledgements

Authors would like to recognize the contributions of Dr. Jane Dennis, University of Bristol; Julia H. Littell, Bryn Mawr College; Elaine Kidney, University of Birmingham; and Library Sciences, University of Oxford.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

References marked with an asterisk indicate studies included in the overview

Achenbach, T. M. (1991). *Manual for the child behavior checklist/4–18 and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.

*Alexander, J. (2004). *Blueprints for violence prevention: Functional family therapy*. Denver, CO. Retrieved from https://www.ncjrs.gov/pdffiles1/ojjdp/204274.pdf

Alexander, J., & Parsons, B. V. (1973). Short term behavioral intervention with delinquent families: Impact on family process and recidivism. *Journal of Abnormal Psychology*, *81*, 219–225.

*Alexander, J., Pugh, C., Parsons, B., & Sexton, T. (2002). *Functional family therapy*. Boulder: Center for the Study and Prevention of Violence, Institute of Behavioral Science, University of Colorado Boulder.

*Alexander, J. F., & Sexton, T. L. (2002). Functional family therapy: A model for treating high-risk, acting-out youth. In F. W. Kaslow (Ed.), *Comprehensive handbook of psychotherapy: Integrative/eclectic* (Vol. *4*, pp. 111–132). Hoboken, NJ: John Wiley.

Alexander, J. F., Waldron, H. B., Robbins, M. S., & Neeb, A. A. (2013). *Functional family therapy for adolescent behavior problems*. Washington, DC: American Association.

*Austin, A. M., Macgowan, M. J., & Wagner, E. F. (2005). Effective family-based interventions for adolescents with substance use problems: A systematic review. *Research on Social Work Practice*, *15*, 67–83. doi:10.1177/1049731504271606

*Baldwin, S. A., Christian, S., Berkeljon, A., & Shadish, W. R. (2012). The effects of family therapies for adolescent delinquency and substance abuse: A meta-analysis. *Journal of Marital and Family Therapy*, *38*, 281–304. doi:10.1111/j.1752-0606.2011.00248.x

Ballard, M., & Montgomery, P. (2017). Risk of bias in overviews of reviews: A scoping review of methodological guidance and four-item checklist. *Research Synthesis Methods*, *8*, 92–108. doi:10.1002/jrsm.1229

Barton, C., Alexander, J. F., Waldron, H., Turner, C. W., & Warburton, J. (1985). Generalizing treatment effects of functional family-therapy—3 replications. *American Journal of Family Therapy*, *13*, 16–26. doi:10.1080/01926188508251260

Becker, L. A., & Oxman, A. D. (2008). Overviews of reviews. In Higgins, J. P. T., & Green, S (Eds.), *Cochrane handbook for systematic reviews of interventions: Cochrane book series*. Chichester, England: John Wiley.

*Bender, K., Tripodi, S. J., Sarteschi, C., & Vaughn, M. G. (2011). A meta-analysis of interventions to reduce cannabis use. *Research on Social Work Practice*, *21*, 153–164. doi:10.1177/1049731510380226

Bornstein, M. H., Hahn, C. S., & Haynes, O. M. (2010) Social competence, externalizing, and internalizing behavioral adjustment from early childhood through early adolescence: Developmental cascades. *Developmental Psychopathology*, *22*, 717–735. doi:10.1017/S0954579410000416

*Breunlin, D. C., Breunlin, C., Kearns, D. L., & Russell, W. P. (1988). A review of the literature on family therapy with adolescents 1979–1987. *Journal of Adolescence*, *11*, 309–334. doi:10.1016/S0140-1971(88)80032-0

Carr, A. (2000). Evidence-based practice in family therapy and systematic consultation: Child-focused problems. *Journal of Family Therapy*, *22*, 29–60. doi:10.1111/1467-6427.00136

Chamberlain, P. R. J. G. (1995). The effectiveness of family therapy in the treatment of adolescents with conduct disorders and delinquency. *Journal of Marital and Family Therapy*, *21*, 441–459.

Costello, E. J., Erkanli, A., & Angold, A. (2006). Is there an epidemic of child or adolescent depression? *Journal of Child Psychology and Psychiatry*, *47*, 1263–1271.

Cottrell, D., & Boston, P. (2002). The effectiveness of systemic family therapy for children and adolescents. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *43*, 573–586.

*Diamond, G. S., Serrano, A. C., Dickey, M. I. T. C., & Sonis, W. A. (1996). Current status of family-based outcome and process research. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*, 6–16.

*Dumas, J. E. (1989). Treating antisocial behavior in children: Child and family approaches. *Clinical Psychology Review*, *9*, 197–222.

Eisner, M. (2009). No effects in independent prevention trials: Can we reject the cynical view? *Journal of Experimental Criminology*, *5*, 163–183.

Elliott, D.S., & Mihalic, S. (2004) Issues in disseminating and replicating effective prevention programs, *Prevention Science*, *5*, 47, doi:10.1023/B:PREV.0000013981.28071.52

*Farrington, D. P., & Brandon, C. W. (2003). Family-based prevention of offending: A meta-analysis. *Australian and New Zealand Journal of Criminology*, *36*, 127–151.

*Filges, T., Andersen, D., & Jørgensen, A. M. K. (2016). Functional family therapy (FFT) for young people in treatment for non-opioid drug use: A systematic review. *Research on Social Work Practice*, *28*, 131–145. doi:10.1177/1049731516629802

Goldstein, A. P., Barry, G., & John, G. C. (1998). *Aggression replacement training: A comprehensive intervention for aggressive youth*. Champaign, IL: Research Press.

*Gordon, D. A. (1995). Functional family therapy for delinquents. In R. R. Ross, D. H. Antonowicz, & G. K. Dhaliwal (Eds.), *Going straight: Effective delinquency prevention and offender rehabilitation* (pp. 163–178). Ottawa, Canada: Air Training.

*Hartnett, D., Carr, A., Hamilton, E., & O'Reilly, G. (2017). The effectiveness of functional family therapy for adolescent behavioural and substance misuse problems: A meta-analysis. *Family Process*, *56*, 607–619. doi:10.1111/famp.12256

Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (2009). *Multisystemic therapy for antisocial behavior in children and adolescents*. New York, NY: Guilford Press.

*Henggeler, S. W., & Sheidow, A. J. (2003). Conduct disorder and delinquency. *Journal of Marital and Family Therapy*, *29*, 505–522. doi:10.1111/j.1752-0606.2003.tb01692.x

*Henggeler, S. W., & Sheidow, A. J. (2012). Empirically supported family-based treatments for conduct disorder and delinquency in adolescents. *Journal of Marital and Family Therapy*, *38*, 30–58. doi:10.1111/j.1752-0606.2011.00244.x

Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration, 2011. Retrieved from www.cochrane-handbook.org

*Hogue, A., & Liddle, H. A. (2009). Family-based treatment for adolescent substance abuse: Controlled trials and new horizons in services research. *Journal of Family Therapy*, *31*, 126–154. doi:10.1111/j.1467-6427.2009.00459.x

*Huey, S. J. J., & Henggeler, S. W. (2001). Effective community-based interventions for antisocial and delinquent adolescents. In J. N. Hughes, A. M. La Greca, & J. C. Conoley (Eds.), *Handbook of psychological services for children and adolescents* (pp. 301–322). New York, NY: Oxford University Press.

Klein, N. C., Alexander, J. F., & Parsons, B. V. (1977). Impact of family systems intervention on recidivism and sibling delinquency: A model of primary prevention and program evaluation. *Journal of Consulting and Clinical Psychology*, *45*, 469–474.

Little, G. L., & Robinson, K. D. (1988). Moral reconation therapy: A systematic step-by-step treatment system for treatment resistant clients. *Psychological Reports*, *62*, 135–151.

McCulloch, A., Wiggins, R. D., Joshi, H. E., & Sachdev, D., (2000). Internalising and externalising children's behaviour problems in Britain and the US: Relationships to family resources. *Children & Society*, *14*, 368–383.

McWey, L. M., Cui, M., & Pazdera, A. L. (2010). Changes in externalizing and internalizing problems of adolescents in foster care. *Journal of Marriage and Family Therapy*, *72*, 1128–1140.

Measuring Recidivism. (n.d.). Retrieved March 2, 2018, from https://www.nij.gov/topics/corrections/recidivism/Pages/measuring.aspx

Montgomery, P., Grant, S., Hopewell, S., Macdonald, G., Moher, D., Michie, S., & Mayo-Wilson, E. (2013a). Protocol for CONSORT-SPI: An extension for social and psychological interventions. *Implementation Science*, *8*, 99. doi:10.1186/1748-5908-8-9

Montgomery, P., Underhill, K., Gardner, F., Operario, D., & Mayo-Wilson, E. (2013b). The Oxford implementation index: A new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of Clinical Epidemiology*, *66*, 874–882. doi:10.1016/j.jclinepi.2013.03.006

Motgomery, P., Grant, S., Mayo-Wilson, E., et al. (2018). Reporting randomised trials of social and psychological interventions: the CONSORT-SPI 2018 Extension. *Trials*, *19*, 407. doi:10.1186/s13063-018-2733-1

Montgomery, P., Underhill, K., Gardner, F., Operario, D., & Mayo-Wilson, E. (2013). The Oxford implementation index: A new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of Clinical Epidemiology*, *66*, 874–882. doi:10.1016/j.jclinepi.2013.03.006

*Mulvey, E. P., Arthur, M. W., & Reppucci, N. D. (1993). The prevention and treatment of juvenile delinquency: A review of the research. *Clinical Psychology Review*, *13*, 133–167. doi:10.1016/0272-7358(93)90038-N

Oxman, A. D., & Guyatt, G. (1991). Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology*, *44*, 1271–1278.

Petterutti, A., Walsh, N., & Valesquez, T. (2009). *The costs of confinement: Why good juvenile justice policies make good fiscal sense*. Retrieved from http://www.justicepolicy.org/research/78

Pieper, D., Antoine, S. L., Neugebauer, E. A., & Eikermann, M. (2014). Up-to-dateness is often neglected in overviews: A systematic review. *Journal of Clinical Epidemiology*, *67*, 1302–1308. doi:10.1016/j.jclinepi.2014.08.008

Pieper, D., Beuchter, R., Jerinic, P., & Eikermann, M. (2012). Overviews of reviews often have limited rigor: A systematic review. *Journal of Clinical Epidemiology*, *65*, 1267–1273. doi:10.1016/j.clinepi.2012.06.015

Prinz, R. J. (1998). Conduct disorders. In S. B. Alan & H. Michel (Eds.), *Comprehensive clinical psychology* (pp. 527–538). Oxford, England: Pergamon Press.

*Robbins, M. S., Alexander, J. F., Turner, C. W., & Hollimon, A. (2016). Evolution of functional family therapy as an evidence-based practice for adolescents with disruptive behavior problems. *Family Process*, *55*, 543–557. doi:10.1111/famp.12230

Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., … Bouter, L. M. (2007). Development of AMSTAR:

A measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, *7*, 10. doi:10.1186/1471-2288-7-10

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMJ*, *340*, c332.

*Sexton, T. L. (2002). Family-based empirically supported interventions. *The Counselling Psychologist*, *30*, 238–261. doi:10.1177/0011000002302003

Sexton, T. L. (2004a). *Implementing functional family therapy: Practical, theoretical, and empirical findings. 17th Annual RTC Conference*. Tampa, FL.

Sexton, T. L. (2004b. *Levels of evidence for the models and mechanisms of therapeutic change in family and couples therapy*. New York, NY: Wiley.

*Sexton, T. L. (2009). Functional family therapy: Traditional theory to evidence-based practice. In J. H. Bray & M. Stanton (Eds.), *The Wiley-Blackwell handbook of family psychology* (pp. 327–340). Hoboken, NJ: Wiley-Blackwell.

*Sexton, T. L., & Alexander, J. F. (2000). Functional family therapy. Family strengthening series. *Juvenile Justice Bulletin*, 9.

Sexton, T. L., Ridley, C. R., & Kleiner, A. J. (2004). Beyond common factors: Multilevel process models of therapeutic change in marriage and family therapy. *Journal of Marital and Family Therapy*, *30*, 131–149. doi:10.1111/j.1752-0606.2004.tb01229.x

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

*Sigrid, J., Qais, A., & Zepeda, V. (2013). Effectiveness and implementation of evidence-based practices in residential care settings. *Child and Youth Services Review*, *35*, 642–656. doi:10.1016/j.childyouth.2013.01.007

*Stanton, M. D., & Shadish, W. R. (1997). Outcome, attrition, and family–couples treatment for drug abuse: A meta-analysis and review of the controlled, comparative studies. *Psychological Bulletin*, *122*, 170–191.

*Tanner-Smith, E. E., Wilson, S. J., & Lipsey, M. W. (2013). The comparative effectiveness of outpatient treatment for adolescent substance abuse: A meta-analysis. *Journal of Substance Abuse Treatment*, *44*, 145–158. doi:10.1016/j.jsat.2012.05.006

*Tarolla, S. M., Wagner, E. F., Rabinowitz, J., & Tubman, J. G. (2002). Understanding and treating juvenile offenders: A review of current knowledge and future directions. *Aggression and Violent Behavior*, *7*, 125–143. doi:10.1016/S1359-1789(00)00041-0

*Tolan, P. H. C. R. E. (1986). Family therapy with juvenile delinquents: A critical review of the literature. *Family Process*, *25*, 619–650.

*Vaughn, M. G., & Howard, M. O. (2004). Adolescent substance abuse treatment: A synthesis of controlled evaluations. *Research on Social Work Practice*, *14*, 325–335. doi:10.1177/1049731504265834

*Waldron, H. B., & Turner, C. W. (2008). Evidence-based psychosocial treatments for adolescent substance abuse. *Journal of Clinical Child and Adolescent Psychology*, *37*, 238–261. doi:10.1080/15374410701820133

Waldron, H. B., Robbins, M. S., Neeb, A. A., & Alexander, J. (2013). *Functional family therapy for adolescent behavior problems*. Washington, DC: American Psychological Association. doi:10.1037/14139-000.

*Weisz, J. R., Hawley, K. M., & Jensen Doss, A. (2004). Empirically tested psychotherapies for youth internalizing and externalizing problems and disorders. *Child and Adolescent Psychiatric Clinics of North America*, *13*, 729–815. doi:10.1016/j.chc.2004.05.006

4 RESULTS OF THE IMPLEMENTATION RESEARCH

The following pages reproduce an article on the Implementation Research, which has been submitted for publication and is currently In Press in Research on Social Work Practice (2019).

Publication title: Implementation Issues in Functional Family Therapy: A Narrative Analysis of the Evidence.

Implementation Issues in Functional Family Therapy: A Narrative Analysis of the Evidence

Clio Weisman

Paul Montgomery

University of Birmingham UK

Abstract

This analysis of Functional Family Therapy (FFT) studies examines whether their variable outcomes are attributable to implementation issues**.** Studies were identified firstly, by way of a recent overview, supplemented by an update of a highly sensitive search including 15 databases, 10 websites, all existing relevant reviews, gray literature as well as contacting experts in the field. Updated searches were conducted in August 2018 and were analyzed according to the Oxford Implementation Index and an assessment of supervision quality. In total, the search yielded 150 records; 48 full texts were retrieved of which 32 were excluded leaving 16 studies containing 5320 unique participants included for analysis. There was no evidence of reported harm. Improved training and supervision were associated with better core outcomes. Although there was no apparent dose relationship**,** it appears that implementation issues are important and also that class and ethnicity were identified for areas of further study.

KEY WORDS

Functional Family Therapy, FFT, delinquency, implementation, juvenile justice, analysis

Implementation issues in Functional Family Therapy: A Narrative Analysis of the Evidence

Currently, Functional Family Therapy (FFT) is indicated as an effective treatment option for behavioral disorders in youth age 10 to 18. While there is a great deal of uptake and support for this intervention, results from systematic reviews, meta analyses and a recent overview suggest considerable heterogeneity in outcomes. The extant literature has indicated that implementation concerns may be the cause for much of the variability across trials and between outcomes, thus prompting this implementation analysis which is based in part on the Oxford Implementation Index.

Behavioral Problems in 10 – 18 year old youth

Behavioral problems refer to a wide array of psychiatric and psycho-social diagnoses, including Oppositional Defiant Disorder (ODD), Attention Deficit Hyperactivity Disorder (ADHD) and substance abuse disorders. Such problems in youth have an estimated prevalence as high as 20% and are often predictors of later symptoms including depression, alcohol and substance use, lower employment rates, delinquency, removal from the home, conduct disorder and criminal actions (Barlow, 2012). This group of disorders affects not only the youth and family but extends to have a negative impact on communities and society at large. Juvenile delinquency is often a result of these behavioral problems and issues that begin in the home or at school often transition into more serious criminal offenses (Thapar et al., 2015). With juvenile delinquency and behavioral disorders continuing to be a significant problem in the United States and worldwide, there must be continued efforts to address the specific needs of disparate client populations. There are a number of approaches to treating troubled youth and their families; the evidence surrounding and supporting these interventions must continue to be built in a high

quality and scientifically sound method. A number of interventions have been designed, developed and researched in the ongoing effort to prevent, reduce and treat these damaging and harmful behaviors. While the research base for these interventions is varied, family-based or systemic approaches may be the most effective (Muncie, 2015; Thapar et al., 2015).

Current treatment approaches

The past twenty years have seen a shift towards adopting research-supported practice and interventions, particularly in the area of family and youth treatment (Littell & Shlonsky, 2009). These often manualized and structured approaches are continuing to gain favor, and there exist a number of interventions that have been presented and accepted as having a strong evidence base which is used to support their continued adoption and uptake. Functional Family Therapy (FFT) is among these EBPs, and its uptake continues to increase. FFT is delivered in over 270 different locations predominately in the United States and United Kingdom; also it is endorsed as being effective by a number of key agencies including Blueprints for Violence Prevention (NCJRS.Gov), Action for Children in the UK (Actionforchildren.org.uk), and the Center for Disease Control and Prevention (CDC, 2018). It is important to examine family approaches and FFT in light of its inherent costliness, the number of individuals and families it serves (estimated to be over 50,000) and possible opportunity costs; it is vital that stakeholders have the best information and evidence around these interventions so their decision making is based on high quality research and reporting. After nearly four decades of use, the implementation of FFT has not been adequately examined or addressed in the research literature. Currently the published implementation studies focus primarily on the engagement process, or, ‘delivery’ of the intervention (Celinska, Furrer, & Cheng, 2013); Hartnett, Carr, Hamilton, & O'Reilly, 2017; McPherson & Macnamara, 2017;). This leaves out a great deal of valuable information regarding

the dose, uptake and context, three implementation elements which may factor considerably in the outcomes of FFT.

What is FFT?

FFT is firmly rooted in systems theory, and while it does share characteristics with other family approaches, it differs in its strict manualisation and phasic approach to treatment. Additionally, FFT is presented as a time-limited treatment, and should not take longer than 3 to 6 months. For further detail relating to the phases and processes of the intervention see Alexander, Waldron, Robbins, & Neeb (2013). FFT reflects a core set of theoretical principles, in which behavior is seen as a representation of the family relational system; i.e., as indicative of the communication, patterns and purposes of the family. The overarching goals of FFT may be described as follows: To change the maladaptive behaviors of youth and families, especially those identified as resistant to change; to reduce the personal, societal, and economic consequences of disruptive behavior disorders; to achieve core outcomes with less cost than many other interventions available. These core outcomes are 1) reduction in recidivism; 2) reduction in substance use and abuse and 3) reduction in out-of home placements. FFT's manualized approach includes reframing, interrupting of negativity or blame, redirection of focus, interpretations of patterns of maladaptive behavior with links to emotions, a deepening understanding of actions, and communication training with focus on positive communication. It incorporates theories of information processing, social cognition, and the psychology of emotion (Alexander et al., 2013). The model includes a Clinical Services System (CSS), an online portal that maintains all session notes, relational assessments, and therapist and client self-reported adherence and fidelity to the model guidelines. This system is monitored by both on and off-site supervisors who are well versed in the model.

Systemic theory posits that a child exists in a series of interconnected systems, Bronfenbrenner's theory of ecological development places a child within five levels of relatedness ranging from a micro system such as the immediate family to the macrosystem which refers to the culture or community in which the child lives (Bronfenbrenner, 1979). Using this foundational theory, and when considering that a child's development and growth often begin and are anchored in the familial relationship, it follows that when treating behavioral difficulties, the family may be at the center of the treatment approach. Systems theory and family approaches to intervention are closely linked and one can be seen as informing the other; the beginnings of familial approaches are found in systems theory (Thapar et al., 2015). The current research base suggests that family interventions may be beneficial to youth aged 10 to 18 presenting with behavioral disorders as well as for those in danger of removal from the home for protective or legal reasons. Family interventions may also reduce the length of time spent by youth in institutions or in care (Woolfenden, Williams, & Peat, 2002).

A systemic approach to treatment will, normally, engage not only with the child, or even with the child and his or her family, but also with the larger systems such as the school, peers, church, probation and parole, child services and other community elements as needed. Systemic interventions are based in large part on understanding the connectedness of all these systems and influencing one while ignoring any of the others will undercut and lessen the value of any intervention or positive change (Thapar et al., 2015).

FFT is a short-term (90-day/8-30 hour), intensive and comprehensive program designed for the treatment of behavioural misconduct in youth aged 10 - 18. It is a treatment modality that can be delivered in clinical settings, school settings, or in the home. It is intended to address those youth and families with a wide array of presenting problems including but not limited to;

criminal behaviour, truancy, running away, sexual misconduct, substance abuse, risk of out of home placement and as a re-entry program for youth returning to the community following release from institutional settings. Recent high-quality evidence has challenged the overall effectiveness of the intervention (Weisman & Montgomery, 2018; Hartnett et al., 2017; Humayun et al., 2017; Darnell & Shuler, 2015). However, some studies suggest that there may be implementation explanations for the heterogeneity of outcome results across studies which is in line with broader psychology trial literature (Cunningham et al., 2018; Gottfredson & Gottfredson, 2002; Hawe, Shiell, & Riley, 2004).

FFT: Current Evidence and implementation

FFT's evidentiary status is inconsistent with systematic review results varying considerably and while it has not been shown to cause harm, there are questions regarding not only its efficacy and effectiveness, but also of implementation and fidelity (Baldwin et al., 2012; Bender et al., 2011; Hartnett et al., 2017; Weisman & Montgomery, 2018). Implementation may be measured in a number of ways, and includes the information related to an intervention's design, delivery, uptake and context. These four domains represent a broad spectrum of implementation elements, and aid in the more complete and thorough understanding of specific interventions; the domains can be adapted and different components or aspects of each may be identified by reviewers as being more germane to their particular topic or issue (Montgomery et al., 2013; Fixsen et al., 2005). It is important to note that harms may need to be considered more widely in order to determine what interventions to implement. While harm refers most commonly to injury, such as poisoning or psychological problems for an individual, it may also include opportunity costs of not spending on interventions with proven effects and thus spending

valuable and limited resources in areas where we are uncertain of the outcomes (Bonell et al., 2012).

Conceptual Basis for Implementation Evaluation

The examination and analysis of implementation is a useful method of understanding the how and why of intervention effects, and for whom it may be effective. As evidence accumulates, there is naturally a progression from *if* a given intervention works to questions around its effectiveness for various populations and implementations- that is, for whom might it work and under what circumstances. It is understood that the transfer of an intervention from controlled clinical trials to broader contexts is dependent, at least partially, upon implementation concerns and processes (Durlak, 1998). The importance of implementation is apparent in three ways: first, it allows for understanding the effects of an intervention in practice; second, it assists in the better understanding of evidence supporting an intervention (mediation and moderation effects); and third, it is needed to facilitate the optimal uptake of evidence-based practices (Durlak & DuPre, 2008; Proctor et al., 2011). Intervention implementation presents a vital piece of how and why differences exist between primary trials. Implementation science, as a field, is broadly concerned with the scientific methodology that will contribute to, and promote, the uptake of quality research and evidence into common practice (Ghate, 2018; Mihalic, 2004). This work contains within it a number of aims, including the translation and uptake of research into practice, obtaining a greater and more complex understanding of what influences the outcomes of an intervention, and what are commonly referred to as evaluation frameworks (Elliot & Mihalic, 2004).

For any intervention to be successful, it must be both feasible and acceptable across different settings (Bonell et al., 2006). Frequently, changes may be made to better suit the

characteristics of the participants or the setting. Different risk factors, prevalence and manifestations of a social problem exist in different cultural and social settings, and thus responses may differ. These are frequently cited as the rationale for adaptation of an intervention (Moore et al., 2006).

Fidelity, broadly, is focused around noting adherence to the proscribed model guidelines and to the intervention manual, and while there are a number of ways to define implementation *fidelity*, the focus has traditionally centered on the need to remain adherent to the core components or active ingredients of an intervention (Mihalic, 2004; Arthur & Blitz, 2000). Strict manualisation has been seen as integral to the effective adoption of complex interventions. The rationale behind this is to keep the techniques, activities, and practices at the core of an intervention intact across different settings. However, it may be determined that adaptation to local circumstances is required which creates a tension between itself and manualisation. To address this issue a new approach and understanding of implementation is emerging where the emphasis shifts from standardization to a consideration of the underlying theories and mechanisms of change (Hawe et al., 2004). A 'theory of change', refers to the underlying principles, logic and rationale that link what an intervention does, and why or how these goals are accomplished. The outcomes that are intended by an intervention must have a clear and delineated course along which users progress through a causal pathway (Bonell et al., 2012). This current and more complex understanding of implementation attempts to understand and focus on the theory of change, suggesting that this will allow for adaptation and the maintenance of effect. What this essentially means is that the function, rather than the activities, may be what is most important in the successful adoption and adaptation of a complex intervention (Bonell et al., 2012). This functional approach does not need to undermine the intervention, rather it allows

for both flexibility and theoretical adherence to the principles underpinning and supporting an intervention, which will in turn produce similar outcomes to other settings.

The role of the developer in interventions has come to the fore in recent research. It seems that larger, more positive effects are found in developer-led trials than in independent studies (Shadish, 2002; Eisner, 2009). There may be alternative explanations for these differences. The 'high fidelity' view posits that the implementation is of better quality *because* the developers are responsible for it. Thus, there is not an issue with the results, but only with their generalizability in a different setting delivered by different people. The 'cynical' view suggests that conflict of interest or biases (which may not be simply financial but also ideological in nature) may explain the differences in results (Eisner, 2009).

Objectives

The purpose of this narrative analysis is to examine the implementation of FFT as it modifies the intervention effects by way of the Oxford Implementation Index (Montgomery, Underhill, Gardner, Operario, & Mayo-Wilson, 2013). That is to say a consideration of the dose, delivery, uptake and context across trials. Further, to better understand how these elements, as well those of supervision quality and amount, may impact the differential effects found. This analysis stems from the published overview of FFT, and the protocol attached to the overview also supports this work (Weisman & Montgomery, 2018).

The core research question therefore is: To what extent are the differential effects found for FFT influenced by implementation factors?

Method

Some researchers have suggested that a high quality implementation of a poor intervention may be more effective than a low-quality implementation of the best practice

intervention (Gottfredson & Gottfredson, 2002). Understanding and analysis of these aspects of FFT research will create a broader and more nuanced understanding of the mechanisms and delivery of FFT across diverse populations. The intervention was designed and trialed primarily in Salt Lake City Utah, with a largely white middle class population. However, given that FFT is now being used with populations and locations that are very different, this is not likely to be the best reference point for its generalizability. It is important to understand and identify the elements of the intervention which are working and result in positive outcomes, and possible areas that require modification or adaptation. Recent overview and meta-analytic evidence illustrates the variability in core outcomes, thus confounding any certainties regarding the implementation and usage of FFT with such a heterogenous population and across varied settings (Weisman & Montgomery, 2018; Hartnett et al., 2017).

Search Strategy for Identification of Studies. All published reviews of FFT which were identified in Weisman and Montgomery (2018) were examined for lists and tables of included studies. Additionally, search terms were modified and used across relevant databases in order to update this overview and capture any relevant studies that were not reported or included elsewhere. Electronic searches were made of relevant databases, government policy reports and professional websites. Experts known to have conducted trials of FFT were contacted in order to augment our search of the grey literature and author called upon personal and professional resources to locate any studies that were complete but perhaps not yet published. There were no publication or geographic limitations although searches were conducted only in English. This search was designed to be highly sensitive and to capture all relevant studies and publications relating to this project. The updated search strategy was conducted in August 2018**.** The search terms (modified as necessary for specific databases) for MEDLINE, PsycINFO, Web of Science

and Cochrane Central Register of Controlled Trials were deliberately sensitive and as follows, adapted for each database as needed:

1. Functional Family Therapy. 2. (trial$ or outcome$ or effect$ or study or analysis or implementation). 3. 1 and 2.

Criteria for considering studies in this analysis.

Study Design. Trials that were focused on randomized controlled trials (RCTs) and quasi-experimental designs (QEDs) which met the Cochrane Effective Practice and Organization of Care Review Group criteria (Higgins & Green, 2002) were included however they were not the sole source of information relating to implementation, adaptation or fidelity details. The inclusion criteria designed for use in the recent overview of FFT was also utilized for this analysis, with necessary additions to capture implementation and process studies (see Weisman & Montgomery, 2018). Still, of the identified trials, many were necessarily excluded because of failure to comply with criteria relating to outcomes (e.g. reporting only length of silences, defensiveness or self-reported family function); not containing a specifically licensed FFT program; or inappropriate referral information of participants. Expansions, interpretations, independent adaptation, or otherwise non-specified FFT models were not included. Functional Family Therapy – Child Welfare (FFT-CW) Functional Family Probation and Parole (FFP) and Functional Family Therapy Gang (FFT-G), which are newly expanded versions with different elements and core components of FFT were not considered for inclusion in this analysis.

Process evaluations and observational studies with no comparator group were examined for implementation and fidelity information. These studies also must have observable outcomes and concrete ratable information relating to the results of the intervention. Complete screening and extraction forms can be supplied upon request.

Types of participants. Youth aged 10 – 18 presenting with behavioral problems including Oppositional Defiant Disorder (ODD), Conduct Disorder (CD), Attention Deficit Hyperactivity Disorder (ADHD), truancy, sexual misconduct, substance abuse and other acting out or delinquent actions. Youth at risk of removal from the home or returning after either placement or incarceration are also included.

***Types of outcome measures reported*.** Implementation data were included and considered with regard to core outcomes of FFT which are as follows: reduction in recidivism rates, fewer and shorter duration in removals from the home, and reduction of substance use or abuse (Alexander et al., 2013). Acceptable objective outcome measures include police and court records, out of home placement or the termination of parental rights, drug screening tests and school reports of truancy, suspensions or expulsions. FFT LLC, the dissemination organization for FFT, propose in their training, media and outreach materials that follow up periods should last anywhere from immediately after the conclusion of treatment to five years post-intervention, and the three primary outcomes of FFT have been used as measurement of the intervention since its original development (Alexander et al., 2013). While the primary outcome measures for FFT are clearly outlined, for purposes of this implementation analysis objective data relating to dose, delivery, adherence, supervision, training of therapists and context were also included and analyzed. The role of the researcher performing the study was also noted, coded appropriately and examined for possible conflict of interest, determining whether they are a developer, designer or stakeholder in FFT, or otherwise have a vested interest in the outcome of the study.

Data extraction and analysis. A data collection form, based on that used in the overview was used and all data coded and put into an extraction sheet. The four domains of dose, delivery, uptake and context which are set forth in Montgomery et al. (2013), were adapted in line with

this framework to be specifically appropriate to this intervention. Supervision amount, quality, and the training of therapists was recorded and narratively synthesized. Numbers of unique individual participants were used to avoid "vote counting" as participant numbers vary considerably across studies. When possible, therapist ratings relating to supervision were informed by CSS data, as well as in accordance with proposed guidelines for the measurement and valuation of therapist supervision relating to psychiatry and psychology (Weerasekera, 2013; Fairburn & Cooper, 2011). All data were double-coded and a third party resolved any disagreements. For additional information regarding data collection contact first author.

Results

Results of the search

The search strategy yielded 150 records of which 102 were excluded at abstract level; 48 full texts were retrieved. These texts reported primary studies and after examination of texts 32 were excluded for failure to meet inclusion criteria; most commonly the study did not contain a specific and licensed FFT program or it was conducted before the formalization of the model (10 studies with n = 419), or outcome measures were not sufficient for inclusion (11 studies with n = 1063). See Table of Included Studies (Table 1) and Table of Excluded Studies (Table 2) in line with Cochrane Handbook methodology below (Higgins & Green, 2011). Finally, 16 trials or studies of FFT were included for analysis in this paper falling into the categories of RCTs (n = 1363), non-random assignment with comparison groups (n = 3795) and those trials with no comparison groups (n = 156). See the PRISMA chart below (Figure 1).

Included studies

Dose. The number of sessions or hours of contact was reported in 8 studies (n = 1235), and ranged from 9-24 sessions, or approximately 30 hours of treatment. FFT guidance requires

12-15 sessions to be completed in approximately 3-6 months. There was no evidence of overall effect found for dose on any significant outcomes (see Table 1). For example, in a single study (n = 135) families received 24 sessions (Friedman, 1989) and reported no significant differences on objective outcomes, but participants did report they felt satisfied with treatment. Darnell and Schuler (2015), a study containing 524 participants, report an average of 9 sessions, which is significantly below the recommended dosage, and results seem broadly in line with studies run by independent evaluators. However, one QED study (n = 187) examined the length of time to arrest as it related to the number of sessions completed. The authors suggest that "those teens who completed only six or less [sic] sessions were significantly more likely to be arrested sooner after intake compared to those attending seven or more sessions" Sholevar, Baron, Aussetts, & Spiga, 2010).

Delivery and Uptake. These domains are essentially reported as number of sessions completed by both therapist and clients. In order to make any determinations specific to delivery and uptake separate from each other, information would need to be provided relating to how many sessions therapists attempted, or how many times therapists arrived for sessions, versus how many times clients did not attend. There are insufficient data reported that is specific to delivery or uptake in any of the included studies to enable this comparison, and there is no indication of variability between number of sessions provided and number of sessions attended.

Context. Across all studies, the racial composition was predominately white, middle class and from low risk environments or neighborhoods. When information was included relating to location, the intervention was most often delivered in the home. 5 RCTs contained 1363 unique participants, of which over two thirds were Caucasian and over half came from two-parent households (Alexander & Parsons, 1973; Friedman, 1989; Humayun et al., 2017; Sexton &

Turner, 2010; Waldron, Slesnick, Brody, Turner, & Peterson, 2001). Those studies with comparators (n = 3795) also contained a largely white participant group. Although one independently led study in this group (n = 187) held 76% African American participants from impoverished, violent and high risk urban areas. This study shows a longer time until subsequent arrest after completion of the program (Sholevar et al., 2010). Darnell and Schuler (2015) contained participants from a largely African American and Hispanic high-risk population. This study (n = 524) reports fewer out of home placements immediately after completion but "by 36 months post treatment all groups had similar out of home placement rates" (p. 78). Across the no-comparator studies, three of the four studies (n = 70) contained a predominately white and middle-class population located in Utah (Alexander, Barton, Schiavo, & Parsons, 1976; Barton, Alexander, Waldron, Turner, & Warburton, 1985; Mas et al., 1985).

Some research has indicated that the therapeutic relationship, specifically the ethnic matching between patient and provider, may be an important factor in positive outcomes (Cabral & Smith, 2011). Flicker et al (2008) was conducted in New Mexico and contained 86 participants, half of whom were Hispanic; in this case, outcomes showed a decrease in substance abuse when Hispanic clients were matched with Hispanic therapists.

A number of studies reported in their discussion that they suspect gender and age of the identified youth may have impacts on the objective outcomes of FFT. These studies also brought attention to possible gender alliances between therapists and identified clients and noted, "different therapist gender conditions create a context that elicits varying responses from both therapists and family members" (Mas et al., 1985, p. 414). Celinska et al (2013) also state that, in their study examining whether FFT participants improve across the seven domains covered in the Strength and Needs Assessment (SNA), the study "uncovered a significant reduction in

emotional and behavioral needs and in risk behavior among participants…Only male participants improved on the Child Strengths Scale"(Celinska et al., p. 32). This may suggest that FFT has a stronger effect in this area on male participants than their female counterparts.

Supervision amount, quality and adherence. Supervision *amounts, quality and adherence* vary greatly in their reporting and analyses across trials of whatever design although no trials were found whose primary aim was to assess the supervision element and the outcomes under study. From the data available, the supervision levels across any trial conducted by developers or stakeholders were rated highly on all measures (Alexander and Parsons, 1973 (n=86); Flicker et al, 2008 (n=86) ; Sexton and Turner, 2010 (n=917); Waldron et al., 2001 (n=114)). In regard to the RCTs, therapists received high amounts of consistent and intensive training, and sessions were supervised rigorously with expert FFT clinicians, with providers receiving between 1.5 to 6 hours of face to face supervision per week. In some studies, sessions were videotaped and rated for adherence to model guidelines and activities (Waldron et al., 2001; Flicker et al., 2008) and CSS was utilized to measure model adherence as well. Additionally, 3 of these 4 studies (n = 286) reported significantly positive effects on objective outcomes. Notably, the most recent of these studies, while being rated as containing high levels of both supervision and therapist training, reported that there were no significant differences between FFT and services as usual in adjudicated recidivism rates, but those therapists who were rated as highly adherent achieved better outcomes, and "results indicate that when practiced with model specific adherence FFT resulted in a significant 34.9% and 30% (respectfully [sic]) reduction in felony and violent crimes and a non-significant, 21.1% reduction in misdemeanor crimes" (Sexton & Turner, 2010, p. 346).

In contrast, the two independently conducted RCTs (n = 246) were rated as having low or unknown rates of supervision and clinical training; model specific training is also not rated highly and was not intensive. The supervision quality, time, and intensity were reported as far less comprehensive or there was no information reported relating to these elements (Friedman, 1989; Humayun et al., 2017). These two independently conducted RCTs found in one case that "no significant difference was found between the two groups in the degrees of improvement reported" (Friedman, 1989, p. 346). The other study "failed to show greater reductions in offending and antisocial behavior in the group allocated to FFT (Humayan et al., 2017, p. 1031). Information was not included in part because Friedman (1989) was conducted prior to CONSORT reporting guidelines. Neither of these independent studies reports significantly positive findings, and there is a lack of data relating to supervision amount and quality, although supervision may be part of the explanation.

Poorer reported supervision and oversight appeared to be a feature of the seven studies with quasi-experimental designs, (n= 3795), five were conducted by independent researchers (n = 3686) and were rated as having either unknown or low levels of supervision amount and quality. There was little information relating to intensity of therapist training with the model or therapist characteristics; supervision quality is not reported comprehensively (Baglivio et al., 2010; Barnoski, 2004; Celinska et al., 2013; Sholevar et al., 2010; Darnell & Schuler, 2015). These studies reported non-significant outcomes on primary outcome measures. Baglivio (2010) reported that they "found few significant differences in the effectiveness of the two modalities" (p. 1050); Barnoski (2004) states "18-month felony recidivism rate for the control group is 27 percent compared with 24 percent for the FFT group" (p. 4).

The two included QEDs (n = 109) conducted by developers of the model were rated as having high levels of supervision, clinical training and oversight, and both reported significantly positive outcome measures (Barton et al., 1985; Robbins et al., 1996). Barton et al. (1985) reported a reduction in recidivism of 33%, which is significantly lower than the rate for the region (63%) and was rated highly on supervision, training and oversight.

The 4 studies with no comparator and classified as observational or process focused contained a total of 156 unique participants. These studies were conducted by developers of FFT and were rated as having high levels of both supervision and therapist training with the model. All four studies reported positive outcomes on not just recidivism rates but also a lessening in defensiveness and increases in alliance (Barton et al., 1985; Alexander et al., 1976).

Therapist training. For the purposes of this analysis, training refers to the education, training, and years of experience of the clinicians in a broad sense, not to FFT-specific training. Providers of the model ranged in experience and training level from undergraduate students with no clinical experience (Barton et al., 1985) to doctoral level clinicians with over 10 years' experience treating individuals and families (Humayan, 2017). However, these matters were inconsistently reported across trials of all designs.

One study (n = 35), reports on an effort to determine whether significant levels of FFT-specific training and oversight would result in positive outcomes when therapy was being conducted by graduate students; therapists received intensive and comprehensive training with the model but had little or no experience as clinicians. Also, every session was supervised in this case. This study did not report significant effects on core outcome measures but did report positive outcomes on reduction of negativity and positive alliance with therapists (Robbins et al.,

1996). This finding is also demonstrated in Barton et al (1985) (n = 27) with the use of undergraduate students undergoing intensive FFT specific training and supervision.

Discussion

Summary of Main Results

This paper analyzed data relating to FFT in order to better understand how the intervention is influenced by implementation and fidelity issues based on an established assessment tool. Across study designs, RCTs, QEDs and observational, studies that are rated highly with regard to not only therapist supervision but specifically the amount and quality of that supervision, resulted in better core outcome results. Importantly, no studies, even those with low or moderate levels of supervisory amount and quality reported any harms which may be a feature of poor reporting practice. However they did *not* show significant effects on core outcome measures. Notably, none of the studies conducted by independent investigators rated highly on quantity or quality of supervisory and oversight measures, which may be attributed to a number of factors.

Potential Bias/Role of the Designers

Of the included studies, approximately half (9 studies containing 1382 participants) were undertaken by trialists who were either original designers and/or developers of the model. All but one of these studies reported positive outcomes (n=1355). Half (7 studies containing 3932 participants) by independent researchers and all but one of these reported null findings (n=3860). Research has shown that there is a possibility of trials and studies conducted by individuals who have an allegiance to the program models they are investigating producing significantly more positive results than those conducted by investigators without such allegiance (Shadish, 2002; Eisner, 2009). Allegiance may be present when even *one* member of a team may benefit, either

financially or otherwise, from a trial's reporting of positive findings (Eisner, 2009). Reported effect sizes of prevention and intervention trials have been shown to be noticeably larger when program developers are involved in contrast to those conducted by independent researchers (Eisner, 2009). The studies included in this implementation analysis demonstrate that those undertaken by the designers of the model do indeed report more significant effects. However, when the issue of implementation and fidelity are also examined, it may not be simply allegiance, but may be also attributable to the fact that the oversight and involvement of intervention designers, the supervision and training of providers, or simply trial size, has a positive influence on the outcomes.

Allegiance effects are not limited to direct fiscal gain of the developer but extends to any member of the research team who may have an interest in the outcomes and may benefit from the findings. The rise of 'best practice' lists may also contribute to biases, in part because these lists aid state and federal agencies in determining what programs will receive funding. This creates a "strong incentive to achieve inclusion on those lists, and there is little doubt that some researchers actively lobby for inclusion of their products in those lists" (Eisner, 2009, p. 171). Additionally, there may be ideological interests which may arise when researchers hold a strong position relating to core issues, and thus may become an advocate for their position that does not allow for them to investigate as objective scientists (Gorman, 2006).

It is important to note that positive effects of an intervention during a study conducted by 'developers-as evaluators' may be attributed to either a) the possibility that the quality of implementation is of a higher standard when it is delivered by a developer of the intervention or b) can be attributed to systemic bias relating to an ongoing conflict of interest by a champion of the program, and possibly related to financial factors (Eisner, 2009). Disentangling which of

these two possibilities is in play becomes paramount when we are attempting to understand what intervention model is most effective and for whom, and perhaps most critically, when we are examining possible harms relating to the uptake of a specified evidence-based model.

Strengths and Limitations of this Analysis

This study analyzed, in a rigorous manner, implementation factors relating to the delivery of FFT. The Oxford Implementation Index provided a useful structure to approach and perform this analysis, and by doing so has begun to clarify some concerning questions regarding the model and its delivery with varied and heterogenous populations.

Analysis of what works for whom is limited because the included studies contain participants who were predominately Caucasian and from middle class families in developed areas of the United States. Of 5314 participants only 111 come from outside the United States (Humayun, 2017), suggesting it may be that FFT implementation has not yet adequately addressed racial and socio-economic issues. The independent researchers flag this issue and pose the question of whether FFT outcomes are influenced by these factors (Celinska et al., 2013; Baglivio, et al., 2010; Barnoski, 2004). However, there is not enough information relating to participant characteristics and outcomes to ascertain the relationship of these issues to the delivery and outcomes of FFT. Also of note is the possibility that gender matching between therapist and youth has a positive influence on the outcomes of FFT, and while research has suggested this as well, there was not adequate data to determine whether this is the case (Wintersteen, Mensinger, & Diamond, 2005).

There have been recent adaptations of the FFT model specific to child welfare (FFT-CW) and youth involved with gangs (FFT-G). There has not yet been enough research reported on these models to include in this analysis, or for a consideration of whether these adaptations were

made with a more functionalist approach to implementation. Until there are more data relating to these adaptations it is not possible to arrive at evidence-informed decisions regarding their use and uptake.

Much of the implementation data were poorly reported, particularly in earlier studies which were conducted before the advent of CONSORT and other reporting guidelines. Future studies should consider using in particular the Tidier Guideline (Hoffman, 2014) and CONSORT- SPI (Montgomery et al., 2018).

Conclusions

This analysis sheds light on whether the effects found for FFT were influenced by differential implementation. Most striking of these is the possible connection between supervision quality and quantity with positive outcomes. Unsurprisingly studies led by the developers of the model rank higher in this area and may indicate that it is the supervision, oversight, and model specific training of therapists that allows them to achieve more success with the model. It may be, however, that it is not merely due to supervision and oversight, but to allegiance effects. The understanding of fidelity is important, as it has been flagged as integral to success of the model; it may not be fidelity in the strictest and most narrow understanding of the word, but a fidelity to the *function* rather than the *form.* It is with continued high quality and in-depth research into the mechanisms of change and implementation of the model, and how FFT is understood, adapted and delivered by therapists to best serve their clients, that we can better understand how to achieve positive effects across an increasingly heterogenous and disparate population of youth and families in need of treatment.

Study	Year	Study Design	Total N	Ethnicity as described	Supervision level	Sessions completed	Location	Allegiance Effects	Significant primary outcomes Yes/No
Alexander & Parsons	1973	RCT	86	mostly white	High	12 to 15	Utah	Yes	Yes
Alexander, Barton, Schiavo & Parsons	1976	OBS	21	mostly white	High	unstated	Utah	Yes	Yes
Baglivio, Jackowski, Greenwald & Wolff	2014	QED	2203	53% white	Low	unstated	Florida	No	No
Barnoski	2004	QED	700	unstated	Low	unstated	Washington	No	Yes
Barton, Alexander, Waldron, Turner, & Warburton; Study 1	1985	OBS	27	mostly white	High	10.3	Utah	Yes	No
Barton et al. Study 3	1985	QED	74	mostly white	Med	15	Utah	Yes	Yes
Celinska, Furrer, & Cheng	2013	QED	72	19% white	High	3.4 months	New Jersey	No	Yes
Darnell & Schuler	2015	QED	524	7.3% white	High	9.1	'Large US City'	No	No
Flicker, Waldron, & Turner, Brody & Hops	2008	OBS	86	50% white	Low	12 to 14	New Mexico	Yes	Yes
Friedman	1989	RCT	135	89% white	Low	24	unknown	No	No
Humayan	2017	RCT	111	90% white	Low	12 to 15	England	No	No
Mas, Alexander & Barton	1985	OBS	22	unstated	High	unstated	Utah	Yes	Yes
Robbins, Alexander, Newell & Turner	1996	QED	35	mostly white	High	unstated	Utah	Yes	Yes
Sexton & Turner	2011	RCT	917	78% white	High	unstated	Utah	Yes	Yes
Sholevar, Baron, Aussetts, & Spiga	2010	QED	187	76% AA	Med	unstated	Philadelphia	No	No
Waldron, Slesnick, Brody, Turner, & Peterson	2001	RCT	114	unstated	High	12	New Mexico	Yes	Yes

Table 1: Table of included studies

Study	Year	Design	N.	Reason for Exclusion
Alexander	1973	OBS	20	Before structuring of FFT, looked at the defensive communication patterns in family therapy recipients
Barton, Alexander, & Turner	1988	OBS	32	Measurement of defensiveness only
Barton et al.	1985	OBS	325	This study focused not on FFT, but whether the training influenced workers decisions to refer youth to placement
Gottfredson, D. C., et al.	2018	QED	129	FFT - G and FFP
Hops, Ozechowski, Waldron, Davis, Turner, Brody & Barrera	2011	RCT	225	Outcomes only related to HIV risk-taking behaviors; FFT integrated with another treatment (CBT)
Klein, Alexander, & Parsons	1977	RCT	86	Same data set as Alexander (1973), but focus on the siblings 30-40 months post treatment
Rohde, Waldron, Turner, Brody, & Jorgensen	2014	RCT	170	FFT in combination with another therapy
Robbins, Alexander, & Turner	2000	QED	37	Measurement of therapist reframing on defensiveness
Robbins, Turner, Alexander & Perez	2003	NR	34	Focus on alliance and dropout rates, not recidivism or other hard outcomes
Slesnick & Prestopnik	2009	RCT	119	In-office therapy conducted with FFT manual as a guide, with no family sessions.
Taxy	2012	RCT		Cost benefit analysis based on Bayesian meta-analysis
White, Frick, Lawing & Bauer	2013	OBS	134	Measurement of callous/unemotional traits only

Table 2: Table of excluded studies

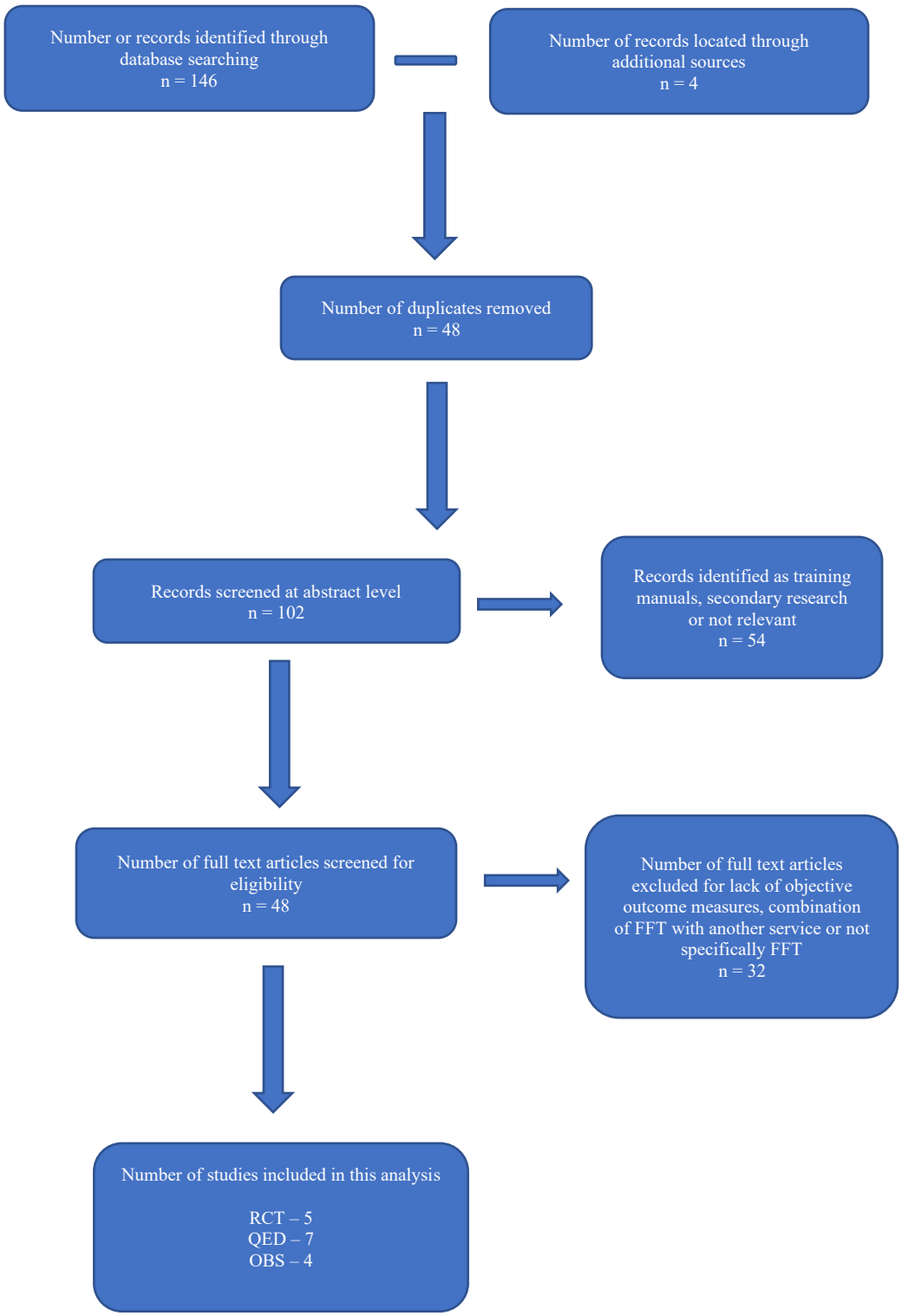

Figure 1: PRISMA diagram

References

Achenbach T. M. (1991). *Manual for the Child Behavior Checklist/4–18 and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Alexander, J. F., Waldron, H. B., Robbins, M. S. & Neeb, A. A. (2013). *Functional Family Therapy for Adolescent Behavior Problems*. Washington, DC: American Association.

*Alexander, J. F. & Parsons, B. V. (1973). Short term behavioral intervention with delinquent families: Impact on family process and recidivism. *Journal of Abnormal Psychology, 81*, 219- 25.

*Alexander, J. F., Barton, C. Schiavo, R. S. & Parsons, B. V. (1976). Behavioral intervention with families of delinquents: Therapist characteristics and outcome. *Journal of Consulting and Clinical Psychology, 44*, 656-64. doi: 10.1037/0022-006X.44.4.656

Arthur, M. W. & Blitz, C. (2000). Bridging the gap between science and practice in drug abuse prevention through needs assessment and strategic community planning. *Journal of Community Psychology, 28*, 241-55. doi: org/10.1002/(SICI)1520-6629(200005)28:3<241::AID-JCOP2>3.0.CO;2-X

Baldwin, S. A., Christian, S., Berkeljon, A. & Shadish, W. R. (2012), The effects of family therapies for adolescent delinquency and substance abuse: A meta-analysis. *Journal of Marital and Family Therapy*, 38: 281–304. doi:10.1111/j.1752-0606.2011.00248.x

Barnett, J. E. (2011). Utilizing technological innovations to enhance psychotherapy supervision, training, and outcomes. *Psychotherapy, 48*, 103–08.

*Baglivio, M. T., Jackowski, K., Greenwald, M. A., & Wolff, K. T. (2014). Comparison of multisystemic therapy and functional family therapy effectiveness: A multiyear statewide propensity score matching analysis of juvenile offenders. *Criminal Justice and Behavior*, *41*, 1033-1056. doi:10.1177/0093854814543272

*Barnoski, R. (2004, March) *Assessing risk for re-offense: Validating the Washington State Juvenile Court Assessment*. Retrieved from http://www.wsipp.wa.gov/ReportFile/868/Wsipp

*Barton, C., Alexander, J. F., Waldron, H., Turner, C. W. & Warburton, J. (1985). Generalizing Treatment Effects of Functional Family-Therapy - 3 Replications. *American Journal of Family Therapy, 13*, 16-26. doi: 10.1080/01926188508251260

Bender, K., Tripodi, S. J., Sarteschi, C. & Vaughn, M.G. (2011). A meta-analysis of interventions to reduce cannabis use. *Research on Social Work Practice*, *21*, 153-64. doi: org/10.1177/1049731510380226

Bonell, C., Oakley, A., Hargreaves, J., Strange, V. & Rees, R. (2006). Assessment of generalisability in trials of health interventions: suggested framework and systematic review. *BMJ, 333*, 346-49. doi.org/10.1136/bmj.333.7563.346

Bonell, C., Jamal, F., Melendez-Torres, G. J. & Cummins, S. (2015). Dark logic: Theorising the harmful consequences of public health interventions, *Journal of Epidemiology and Community Health, 69*, 95.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA; Harvard University Press.

Cabral, R. R. & Smith, T. B. (2011). Racial/ethnic matching of clients and therapists in

mental health services: A meta-analytic review of preferences, perceptions, and outcomes. *Journal of Counseling Psychology,* 58, 537-554. doi: 10.1037/a0025266.

*Celinska, K., Furrer, S. & Cheng, C. C. (2013). An outcome-based evaluation of functional family therapy for youth with behavioral problems. *OJJDP Journal of Juvenile Justice*, *2*, 23-36.

Cunningham, P. B. et al. (2018). Midtreatment problems implementing evidence-based interventions in community settings. *Family Process, 58*, 287-04. doi.org/10.1111/famp.12380

*Darnell, A. J. & Schuler, M. S. (2015). Quasi-experimental study of Functional Family Therapy effectiveness for juvenile justice aftercare in a racially and ethnically diverse community sample. *Child Youth Services Review, 50*, 75-82. doi: 10.1016/j.childyouth.2015.01.013

Durlak, J. A. (1998). Why program implementation is important. *Journal of Prevention and Intervention in the Community, 17*, 5-18.

Durlak, J. A. & DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Clinical Psychology, 41,* 327-50, doi: org/10.1007/s10464-008-9165-0

Eisner, M. P. (2009). No effects in independent prevention trials: Can we reject the cynical view? *Journal of Experimental Criminology, 5*, 163-83, doi: 10.1007/s11292-009-9071-y

Fairburn, C. G. & Cooper, Z. (2011). Therapist competence, therapy quality, and

therapist training. *Behaviour Research and Therapy, 49*, 373-78, doi: 10.1016/j.brat.2011.03.005

*Flicker, S. M., Waldron H. B., Turner, C. W., Brody, J. L. & Hops, H. (2008). Ethnic matching and treatment outcome with Hispanic and Anglo substance-abusing adolescents in family therapy. *Journal of Family Psychology, 22*, 439–47.

Fixsen, D.L., Naoom, S.F., Blase, K.A., Friedman, R.M., & Wallace, F. *Implementation Research: A synthesis of the literature* (2011). National Implementation Research Network, University of South Florida, Tampa, FL.

*Friedman, A. S. (1989). Family therapy vs. parent groups: Effects on adolescent drug abusers. *American Journal of Family Therapy, 17*, 335-47.

Ghate, D. (2018). Developing theories of change for social programmes: co-producing evidence-supported quality improvement. *Palgrave Communications, 4.* doi:10.1057/s41599-018-0139-z

Gottfredson, D. C., Kearly, B., Thornberry, T. P., Slothower, M., Devlin, D. & Fader, J. J. (2018). *Scaling-up evidence-based programs using a public funding stream: A randomized trial of Functional Family Therapy for court involved youth*. Paper submitted for publication.

Gottfredson, D. C. & Gottfredson, G. D. (2002). Quality of school-based prevention programs: results from a national survey, *Journal of Research in Crime and Delinquency, 39*, 3-35. doi: 10.1177/002242780203900101

Hartnett, D., Carr, A., Hamilton, E. & O'Reilly, G. (2017), The Effectiveness of Functional Family Therapy for Adolescent Behavioral and Substance Misuse Problems: A Meta-Analysis. *Family Process,* 56, 607–19. doi:10.1111/famp.12256

Hawe, P., Shiell, A. & Riley, T. (2004). Complex interventions: How "out of control" can a randomised controlled trial be? *BMJ, 328*, 1561-63. doi:10.1136/bmj.328.7455.1561

Higgins, J. P. T. & Green, S. (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org.

Hoffman, T. C. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *The BMJ, 348*, g1687. doi: 10.1136/bmj.g168

*Humayun, S., Herlitz, L., Chesnokov, M., Doolan, M., Landau, S. & Scott, S. (2017). Randomized controlled trial of Functional Family Therapy for offending and antisocial behavior in UK youth. *Journal of Child Psychology and Psychiatry, 58*, 1023-32. doi:10.1111/jcpp.12743

Littell, J. H. & Shlonsky, A. (2009) Towards evidence informed policy and practice in child welfare. *Research on Social Work Practice, 70*, 723-25. doi: 10.1177/1049731509347886

*Mas, C. H., Alexander, J. F. & Barton, C. (1985). Modes of expression in family therapy: a process study of roles and gender. *Journal of Marital and Family Therapy, 11*, 411-15.

McPherson, L. & Macnamara, N. (2017). *Supervising child protection practice: What works? An evidence informed approach*, Springer International Publishing. doi: 10.1007/978-3-319-5003

Mihalic, S. (2004). The importance of implementation fidelity, *Emotional & Behavioral*

Disorders in Youth, 4, 83-105.

Montgomery, P., Underhill, K., Gardner, F., Operario, D. & Mayo-Wilson, E. (2013). The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of Clinical Epidemiology*, *66*, 874-82. doi:10.1016/j.jclinepi.2013.03.006

Montgomery, P., Grant, S., Mayo-Wilson, E., Macdonald, G., Michie, S., Hopewell, S., & Moher, D., on behalf of the CONSORT-SPI Group. (2018). Reporting randomised trials of social and psychological interventions: the CONSORT-SPI 2018 Extension. *Trials.* 19:407. 1-14.

Muncie, J. (2015). *Youth and Crime*. London, UK: Sage.

Oxman, A. D., & Guyatt, G. (1991). Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology*, 44, 1271-78.

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P…& Hensley, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health, 38,* 65-76. doi: 10.1007/s10488-010-0319-7

Robbins, M. S., Alexander, J. F., & Turner, C. W. (2000). Disrupting defensive family interactions in family therapy with delinquent adolescents. *Journal of Family Psychology, 14*, 688–701. https://doi.org/10.1037/0893- 3200.14.4.688

*Robbins, M. S., Alexander, J. F., Newell, R. M., & Turner, C. W. (1996). The immediate effect of reframing on client attitude in family therapy. *Journal of Family Psychology, 10*, 28-34. http://dx.doi.org/10.1037/0893-3200.10.1.28

*Sexton, T. & Turner, C. W. (2010). The effectiveness of Functional Family Therapy for youth with behavioral problems in a community practice setting. *Journal of Family Psychology, 24*, 339-48. doi: 10.1037/a0019406.

*Sholevar, E., Baron, D. A., Aussetts, M. A. & Spiga, R. (2010). Treatment adherence in family therapy with youth as a predictor of time to arrest. *The American Journal of Family Therapy, 38*, 1-15. doi: 10.1080/01926180902754695

Thapar, A., Pine, D. S., Leckman, J. F., Scott, S., Snowling, M. J., & Taylor, E. (Eds.) (2015). *Rutter's child and adolescent psychiatry* (6th edition). West Sussex, UK: John Wiley & Sons Ltd.

*Waldron, H. B., Slesnick, N., Brody, J. L., Turner, C. W., & Peterson, T. R. (2001). Treatment outcomes for adolescent sub- stance abuse at 4- and 7-month assessments. *Journal of Consulting and Clinical Psychology, 69,* 802–813

Weerasekera, P. (2013). The state of psychotherapy supervision: recommendations for future training. *International Review of Psychiatry*, 25, 255-264. doi: 10.3109/09540261.2013.769431

Weisman, C. B., & Montgomery, P. (2018). Functional Family Therapy (FFT) for behavior disordered youth Aged 10–18: An overview of reviews. *Research on Social Work Practice, 29*, 333-46. doi.org/10.1177/1049731518792309

Wintersteen, M. B., Mensinger, J. L., & Diamond, G. S. (2005). Do gender and racial differences between patient and therapist affect therapeutic alliance and treatment retention in adolescents? *Professional Psychology: Research and Practice, 36*, 400-08.

Woolfenden, S., Williams, K., & Peat, J. (2002). Family and parenting interventions for conduct disorder and delinquency: a meta-analysis of randomised controlled trials, *Archives of Disease in Childhood, 86*, 251-56. doi: 10.1136/adc.86.4.251

* Indicates inclusion in analysis

5 DISCUSSION

5.1 Introduction

This thesis contributes to the comprehensiveness and quality of evidence relating to Functional Family Therapy (FFT) and its implementation; this work also aims to contribute to the field through the presentation of a more complete and unbiased picture of this evidence, which will aid policy makers and practitioners in their decision-making process relating to its delivery and uptake. It focused first on the completion of an overview of FFT reviews, and then on a secondary implementation analysis drawn from data reported in primary trials. This chapter contextualizes and synthesizes the prior chapters as well as addressing the concerns and problematic areas which emerged regarding the intervention. The chapter begins with a summary of main results and goes on to address the results of the overview and implementation analysis in greater detail; the elements laid out in the background section of this thesis are addressed specifically as they were found to relate to FFT; allegiance effects and bias are then discussed, followed by a discussion of harms. There is then a reflexive section relating to the processes, personal experience and issues that resulted from the publication of the overview and over the course of this project, and how these difficulties were addressed. Finally, it concludes with addressing the implications for future research and practice.

5.2 Summary of Main Results

Through systematic overviewing methodology and a structured implementation analysis this thesis found that the research base of FFT does not present an adequate, unbiased or complete picture of the evidence. The overview illustrated that the majority of reviews reporting on FFT outcomes are neither reliable nor of a rigorous or high quality. The results reported therein are inconsistent and incomplete and cannot be determined as sufficient to answer definitively whether FFT is effective, nor for whom or under what circumstances this may or

may not be the case. The implementation analysis, which was conducted in order to better understand the mechanisms and implementation concerns which were identified through the overview, demonstrated that FFT outcomes are greatly influenced by implementation elements, some of which may be a result of contextual differences. Additionally, the analysis drew attention to the role allegiance biases play in the reporting and evaluation of the model.

5.3 Findings from the Overview

The overview of FFT clearly indicates that the research base for the intervention is not sufficient and is in fact quite narrow in scope and applicability. Since the advent of FFT in the late 1970's youth criminality has continued to be a persistent and important problem in the United States, United Kingdom and globally. While it is true that since 1980 there has been an overall drop in the number of reported crimes committed by juveniles, this by no means indicates that the pervasive problem is not severe. The graph below indicates that while this drop in crime rates is evident, the fact remains that there are estimated to be over 800,000 juvenile arrests occurring in 2017 in the United States, with over 200,000 of these involving violence against another person.

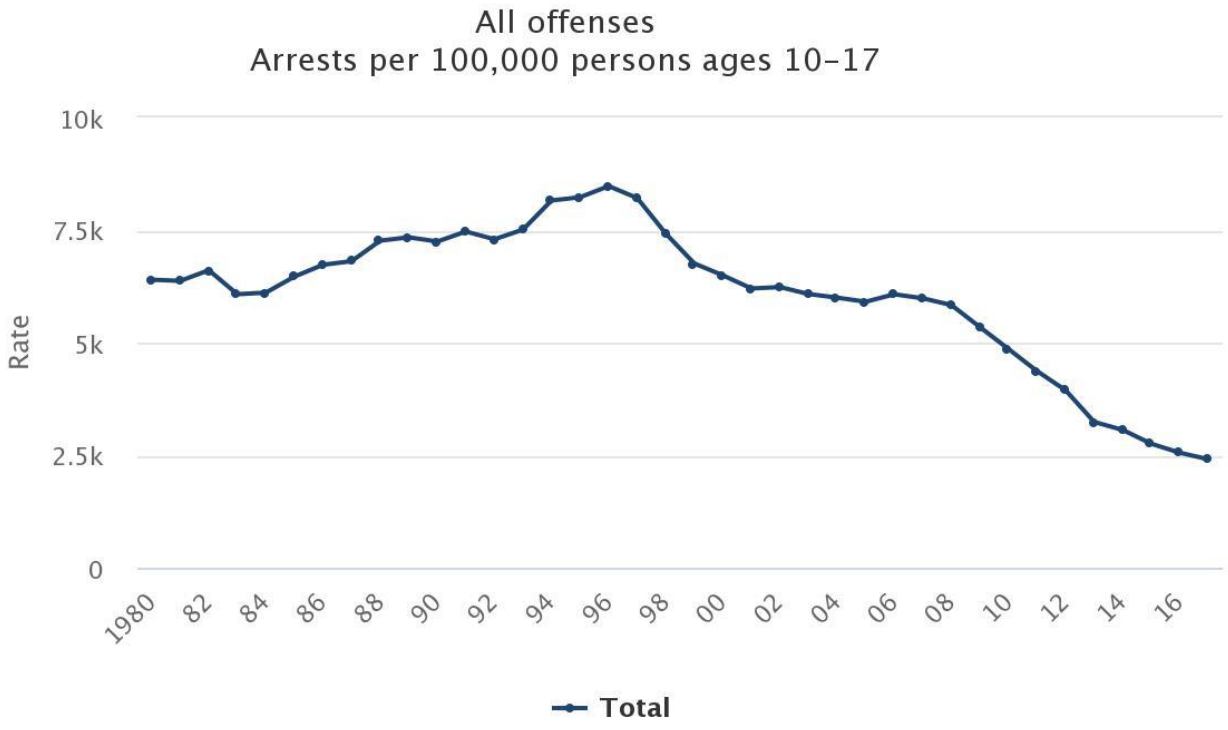

Figure 3. Graph of US juvenile arrest rates 1970 – 2017 (Source: United States Office of Juvenile Justice and Delinquency Prevention; n.d.)

In the United Kingdom, the other main setting in which FFT is continuing to increase in uptake, juvenile crime persists as a significant problem; while records do indicate a decrease in arrests, this does not imply that there is no longer a large-scale problem relating to delinquency and criminality of minors. In 2018 (the most recent year for which data are available) there were approximately 75,000 arrests in England and Wales of persons age 10 – 17 (gov.uk).

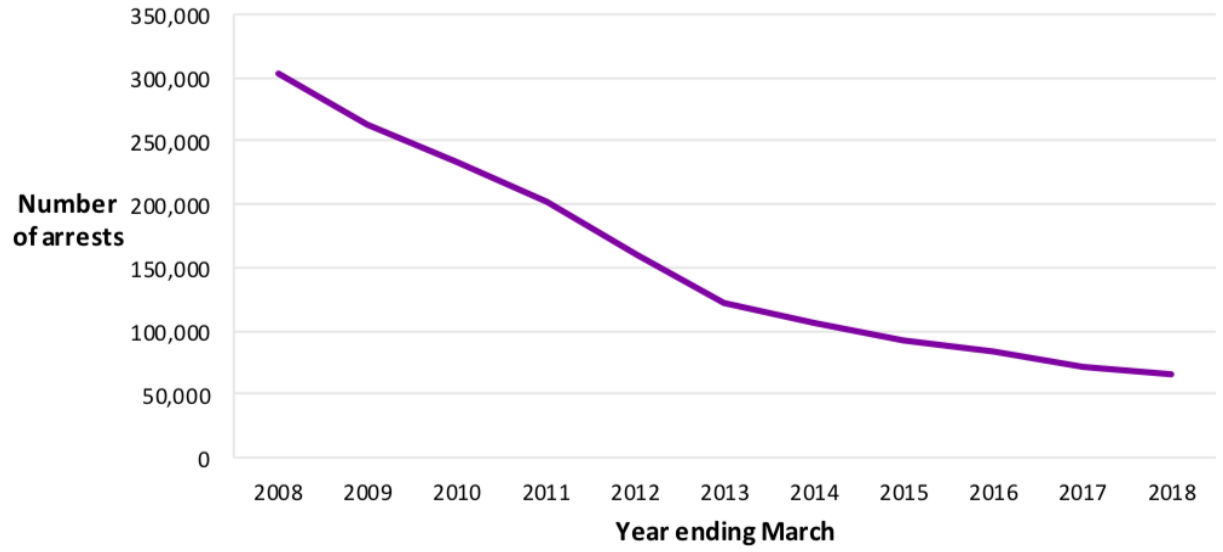

Figure 4. Trends in arrests of children for notifiable offences, England and Wales, years ending March 2008 to 2018 (Source: United Kingdom Ministry of Justice, n.d)

While these issues are also significant problems in developing or low to middle income countries, FFT has not, to date, been taken up in these regions and there is no evidence currently that suggests it would or would not be an appropriate intervention in these settings. The overview contained within it no review that referred to any data found in these contexts and did not indicate that any research has been conducted in this area.

In the face of these data, the overview becomes even more relevant, as FFT is designed specifically to combat this problem. Rather than supporting FFT's claim that it is an effective and evidence supported intervention for the reduction of juvenile delinquency across heterogenous populations and contexts, the overview illuminated the paucity of high-quality and

reliable evidence and reporting while simultaneously drawing attention to the fact that it is likely differences in implementation that accounts for this variability in outcomes. Additionally, the overview drew attention to possible selective outcome reporting and publication bias and determined that the existing reviews of FFT were not an adequate source of determinative information regarding the use of FFT.

Overviews, by definition, are limited to reporting and synthesising only the evidence presented by the individual reviews contained within them, and the quality of these reviews will also be an important factor in the overall completeness of the overview (Ballard & Montgomery, 2017). As discussed in sections 2.2–2.4 of this thesis, overview methodology is useful when attempting to reduce bias and to present a coherent picture of a varied evidence base. The overview of FFT included 31 reviews for analysis, and revealed that, when examined as a whole, the published evidentiary base of FFT is inconsistent with main effects on core outcomes being modest. The variability between outcomes extends from the reviews to the trials included within them, and it became apparent that those reviews authored by designers or developers of the model reported far more significant positive findings. The quality of reviews was also widely variable, with few reviews including risk of bias tables, peer reviewed or critical appraisal methods.

In many ways the overview created more questions than it answered, in part because it was conducted to a higher standard than the majority of available reviews. This standard was accomplished through the use of a pre-determined protocol, highly sensitive and exhaustive search strategy, double-coding, and a structure informed by and supported by Cochrane's guidelines for overview methodology. Despite this, the questions raised were critical to furthering an evidence base free from bias, incompleteness or lack of rigour. The overview also

acknowledged that there must be a continued investigation into FFT, and highlighted the areas of particular concern. The more broad and inclusive design of this overview was not able or intended to pick up the narrow and specific elements of implementation concerns. These questions were drawn out from the overview and addressed through the implementation analysis that followed.

5.4 Results of the Implementation Analysis

The implementation analysis contained within this work is modelled on the Oxford Implementation Index, which examines the domains of dose, delivery, uptake and context and should be altered and adapted to be germane to the particular topic and area of study (Montgomery et al., 2013). It should be noted that the index is only as useful as the information provided by trialists relating to implementation elements. This index was modified to be suitable for the analysis of FFT. Primary trials and studies of FFT were coded according to the following domains:

1) Dose
2) Delivery and uptake
3) Context
4) Supervision amount, quality, and adherence
5) Therapist training
6) Potential bias or role of designers

The analysis showed that there appeared to be no significant or measurable relationship between dose (which was defined as the number of sessions or hours spent in direct treatment) and outcomes, as in all but two studies that reported on this measure the dosage – or number of sessions – was consistently within the range suggested by FFT training materials. FFT is intended to be a time-limited intervention and should not extend too far beyond what is intended

by the model designers. This position and finding reflects what current research suggests, which is that longer treatment does not necessarily mean more effective treatment. This is particularly true in the case of CBT and systemic forms of family therapy, which are both shown to be more effective when they are time-limited (Thapar et al., 2015).

The delivery and uptake domains of FFT studies could not be adequately examined. The delivery and uptake would in the context of FFT be measured through noting the number of sessions therapists *attempted* versus the number of sessions the family was present for and engaged with. In essence, there was no way of measuring the difference, if indeed there was one, between what was presented to the recipients of FFT, and what was absorbed or taken up by them. It may be that this domain is less applicable to psychosocial or behavioral interventions than it is to pharmacological or otherwise more easily quantifiable trials.

The context of FFT trials placed the majority of research in white, middle to upper middle-class urban environments in the United States. This emerged as problematic, in large part because the intervention is said to be culturally sensitive and appropriate across differing racial, ethnic, socio-political and cultural landscapes. The implementation analysis shed light on this particular area, and it became evident that the research into FFT was skewed in the direction of studies containing predominately white participants.

While many of the included studies did not specifically report supervision amount, quality, or adherence, those that did provided support for the hypothesis that increased levels of high-quality supervision and oversight were related to more positive core outcomes. There were not enough data to prove a correlative relationship, but such a relationship is certainly strongly suggested by the analysis.

Therapist training and clinical experience, irrespective of any model-specific training, could not be shown as a demonstrable area of impact on core outcomes, as there were not enough data to support and conclusions regarding connection between more (or less) experienced therapists achieving better outcomes.

The likelihood of allegiance bias in FFT research and literature was hypothesized, and this was borne out in the implementation analysis, which revealed that over half of the included studies were conducted by trialists who were either original designers and/or developers of the model and furthermore that all but one of these trials reported significant positive outcomes across every measure reported; conversely, all but one of the independent studies reported null findings.

The implementation analysis of FFT supported the strong possibility that it may very well be implementation elements that are central to the success of the model. What emerged was a picture of a manualized intervention that was reported as effective almost exclusively when delivered with specific populations, in very narrow contexts and with the oversight and involvement of the original developers and designers of the model. The strong efficacy reported in designer-led trials in white middle class settings may not be replicable in different contexts. This lack of generalisability is of particular interest when considering two main issues. The first is that juveniles from minority and lower income families are far more likely to be involved with the juvenile justice system, and thus more likely to be in mandated treatment programmes such as FFT. African American youth comprise approximately 16% of youth nationwide, yet as of 2018, 44% of the youth being held in juvenile facilities (including detention centres, training schools and juvenile jails or prisons) were black. For more details relating to arrest records and minority over representation see Figure 5 below.

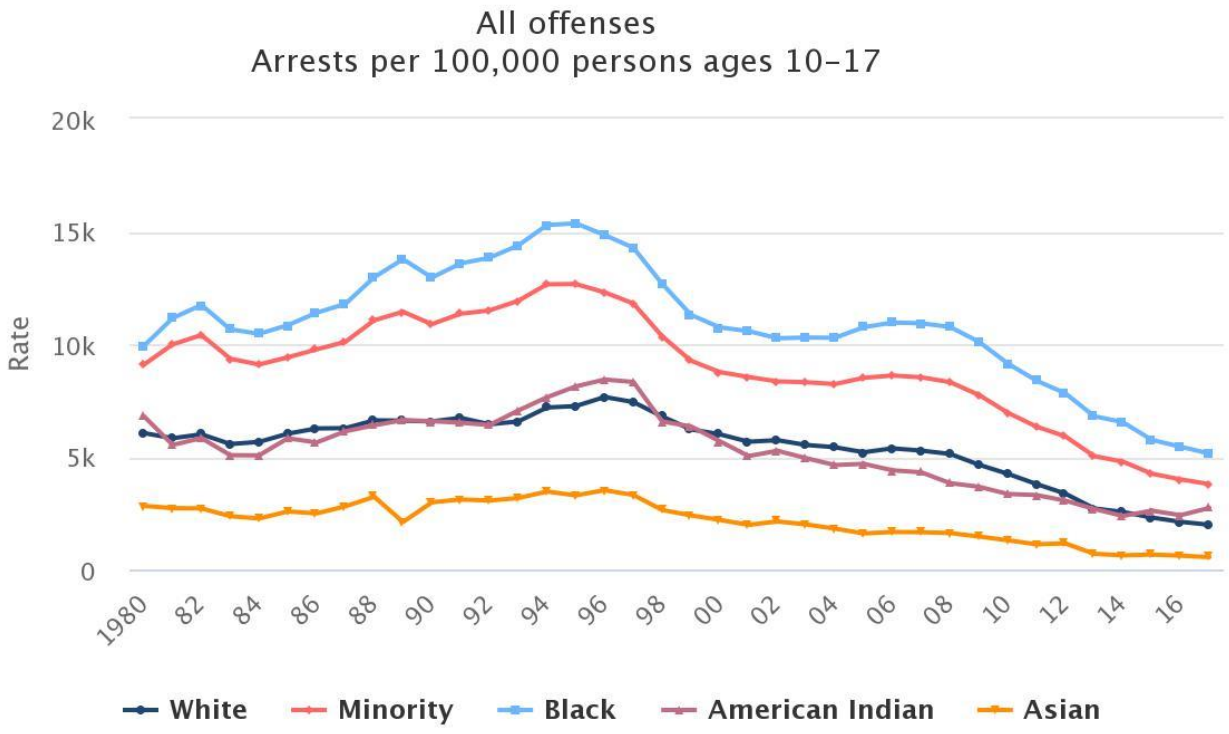

Figure 5. Graph of US juvenile arrest rates by ethnicity 1970 – 2017 (Source: United States Office of Juvenile Justice and Delinquency Prevention, n.d.)

The second issue to keep in mind is that FFT is being increasingly delivered in lower income or minority contexts. It is no longer being delivered to predominately white and middle class youth and their families, but has been taken up in impoverished, ethnically diverse and high crime cities in the United States such as the Bronx, Brooklyn, New Orleans, New Haven, Watts and South Central Los Angeles. These two issues present a serious problem, as the disseminators of FFT are presenting the intervention as effective with these groups and in these contexts without the substantive and definitive evidence to support their claims.

Also important to keep in mind is that implementation fidelity and contextual adaptations need not be in conflict. It is necessary to implement interventions with a consistent theoretical adherence, and perhaps not with specific and *formal* adherence. If the underpinning mechanisms

for change are respected and allowed to frame and support the implementation, it may very well be that there will be successful outcomes as a result. Standardizing an intervention, and allowing for no flexibility whatsoever, will not be the best approach, if the goals are increased uptake and positive outcomes in real world settings. The implementation analysis contained in this thesis supported this position, and contributed to the evidence supporting a functional, rather than formal, approach to implementation and delivery of treatment models.

5.5 Behavioral and Psychological Difficulties in Children and Youth

5.5.1 Childhood/Familial Risk Factors

It is known that stressors such as parental psychosocial or psychological issues will likely contribute to dysfunction, and these risk factors are not adequately addressed by the FFT model (Alexander, 2013; Bowlby, 1969) There is no indication in the materials presented by FFT that there is specified methods designed to address the possibly extremely negative impact of a parental diagnosis or psychological issue (Alexander, 2013; Hartnett et al., 2017). Neither the overview nor the implementation unearthed any evidence that FFT has within it a research supported means of approaching and managing significant familial risk factors such as a pre-existing and deeply insecure attachment style or maternal diagnoses.

5.5.2 Environmental Risk Factors

Strategic and structural therapy models, both of which are contributory to the design of FFT, demand that when treating a juvenile or family, the entirety of the context must be accounted for and addressed. While FFT literature does state this as a goal and tenet of the intervention, in actuality this is not the case. If FFT was truly, as stated in the presenting literature and training materials, culturally sensitive and appropriate to different families and their needs, it would perhaps be more likely to result in positive outcomes in settings so different from where it originated and was predominately trialed (Hartnett et al., 2017). Both the overview and implementation analysis support this supposition.

5.5.3 Diagnoses

When a child or adolescent carries the diagnoses of Oppositional Defiant Disorder (ODD), Conduct Disorder (CD), or Attention Deficit Hyperactivity Disorder (ADHD), he or she is more likely to behave in a manner that results in a referral for treatment. ODD and CD both by definition contain some element of opposition to authority, loss of temper, outbursts, vindictiveness, physical fights, destruction of property, and other externalizing actions, in ODD these will be less extreme, but this disregard for rules and acceptable behaviour will very likely attract the attention of an authority figure. A diagnosis of ADHD indicates that the youth may have trouble focusing, sitting still, controlling impulsivity and exhibition of a low threshold for frustration. Particularly in the case of ODD and CD, there is very likely a criminal aspect to the behaviors. If left untreated, these diagnoses can evolve and progress in severity into increasingly harmful examples of externalizing and acting-out behaviours. These diagnoses do not exist in a vacuum, and both ODD and CD note in their diagnostic criteria that overly harsh, neglectful or inconsistent parenting are commonly found in the families of those carrying these diagnoses (DSM-V). One of the issues that came to the fore through the completion of this thesis is that FFT does indeed approach the entire family system, and not solely the identified patient, which would be an indicated approach when attempting to alleviate the symptom picture of these diagnoses. This is a research supported approach and it must be noted that FFT maintains a position of treating the family, not the individual.

5.6 Reason for referral

FFT is presented as an appropriate course of treatment for a wide array of presenting problems. As noted in the background section of this thesis, it has been recommended by best practice lists and treatment guidelines not just for the treatment and prevention of juvenile delinquency, but also for the reduction of substance abuse, risk of out-of-home placement, sexual

misconduct, and a broad range of familial issues including those families in which abuse has been indicated (blueprintsprograms.org; OJJDP.gov; WHO.int). FFT is recommended by its advocates across a vast range of settings and for any ethnic, cultural or sociopolitical group (Alexander, 2013; Sexton, 2002). When any intervention is declared as an effective intervention for such an exhaustive list of diagnoses and concerns and as useful and appropriate for any and all different populations it is impossible not to speculate as to whether there is rigorous evedence that this is in fact the case. For an intervention to be effective, it needs to be not only shown and demonstrated to be effective in ideal conditions or settings, but acceptable to those to whom it is being delivered. While this is a core element of successful implementation, it is important to note that the reason for referral to a course of FFT must also be considered, as it may very well impact the approach to treatment. In fact, the approach and elements should perhaps be altered, or tweaked slightly, to better match to the specific needs of individuals and communities. While FFT literature states that it is designed for precisely these issues, the real-world delivery of the model may not support this.

5.6.1 Anti-social Behaviours

Those youth who are referred to treatment for antisocial behaviour represent the most noticeable and easily identifiable group of youth in need of treatment or services. These behaviours often result in direct harm to another person, and thus come to the attention of authorities or other referral sources with a greater frequency. It must always be noted that anti-social behaviours, if not addressed appropriately, may often lead to later criminality and more severe criminal outcomes (Muncie, 2015; Thapar et al., 2015). Antisocial behaviours often result in a youth being caught up in the legal system, where FFT has been supported by prior reviews and research as an intervention that will definitively reduce the incidence of criminal, acting out behaviours, and recidivism, but this claim was not found to be entirely supported by the research.

Additionally, the overview and implementation analysis both draw attention to the fact that recidivism is measured and reported differently in different trials and settings. For further detail relating to the measurement of recidivism, refer to section 2.6.2 of this thesis. This is a hugely confounding factor and makes it difficult to draw any rigorous parallels or conclusions. There are no standardized measures for recidivism in place, and the measurement of it varies greatly across trials and dependent upon municipality or region. Recidivism of criminality or antisocial actions are presented as an objective outcome measure, but this is not an accurate portrayal of the evidence. Recidivism may be identified in a number of different ways and there is no truly reliable, consistent or in some cases even accurate means of measurement. For example, agencies in both the United States and the United Kingdom draw attention to the fact that recidivism rates are likely much higher than what is reported due to inaccuracies and differing definitions (www.gov.uk; www.nij.gov). Additionally, there is little to no transparency around how trialists choose to measure or quantify recidivism rates for the purposes of FFT outcomes. The result of these differences is that when a reduction in recidivism is presented as an outcome of FFT, there is often no real way of ascertaining what this entails and how the rates were calculated.

5.6.2 School-related Behaviour

The trials of FFT that were included in the implementation analysis, and the evidence presented by the overview, did not support FFT as an effective intervention for school related behaviours. FFT LLC continues to represent the intervention as effective for this issue, but at this time there is not adequate evidence to support this claim.

5.6.3 Sexual Misconduct

At this time there is no conclusive evidence that a course of FFT will be effective in the treatment of sexual misconduct specifically. The overview did not discover any evidence relating to this behaviour or reason for referral, nor did the implementation analysis find any indication

that FFT may be an appropriate or successful intervention if and when this is the presenting problem. It should be noted that sexual misconduct should not, under almost every circumstance, be treated by anything less than direct and individual treatment (Thapar et al.,2015; Walker, McGovern, Poey, & Otis, 2004). FFT LLC promotes the use of FFT for this type of identified problem, but this is not a research-supported conclusion. Once again, it must be reiterated that this thesis is not intended to determine whether FFT is effective when applied to *specific* referral reasons or identified problems, but one may extrapolate that, given the inconsistency and inconclusive nature of much of the evidence, recommending FFT in response to sexual misconduct is in conflict with current treatment guidelines which recommend intensive cognitive behavioural interventions that target specified risk factors in an effort to reduce the likelihood of reoffending or escalation of offences (Beech et al., 2007).

5.6.4 Indicators of Abuse

When abuse is indicated, it is important, first and foremost, to ensure the safety of all family members. Abuse is very often coupled with family dysfunction. It follows that if that dysfunction is addressed, the abuse will be minimized. Strategic family therapy, upon which FFT is based, might indicate, quite accurately, that certain types of abuse are in actuality malformed efforts at communication. FFT incorporates many elements from strategic family therapy, and does identify abuse as a pattern of behaviour that can be effectively treated through FFT. In theory, helping the family to achieve a higher level of functioning, strengthening protective factors and reducing risk factors, incidents of abuse and interpersonal violence within the home will be minimized. The overview and implementation revealed no evidence in support of this, and given the safety concerns relating to this, promoting the use of FFT without additional supports and treatment when the referral stems from suspected or confirmed abuse is not an appropriate clinical or policy decision.

5.7 FFT in conjunction with other approaches

5.7.1 Pharmacological

Research has suggested that a pharmacological approach is not, on its own, a lasting or effective treatment for behavioral disorders in youth (Loy et al., 2012). However, combining this with additional treatment, such as CBT or a systemic family therapy, may be supported by some research. This approach may be necessary particularly in cases where the aggressive behaviours are so severe as to present a clear and immediate danger to the individual or others (Pappadopulos et al., 2006). FFT may in some cases be recommended or mandated in addition to medication, but there is no evidence supporting this, nor is there any evidence that the effects of FFT will be supported or strengthened by combining it with medication. It may very well be that anti-depressants or anti-anxiety medication assists in the management of symptoms, which in turn allows for increased engagement in family therapy, but neither the overview nor the implementation analysis discovered any evidentiary support for this, and there does not appear to be any research examining this possibility.

5.7.2 Psychodynamic

This thesis did not discover any additional evidence relating to the use of individual psychodynamic therapy or treatment as pertaining to juvenile behavioral problems and delinquency. FFT has not been studied in combination with individual psychodynamic therapy. Still, if a presenting youth has significant psychological problems that are more internalizing and emotional in nature than externally disruptive, it may be that individual psychotherapy would be recommended and beneficial for the management of his or her symptoms (Fonagy, 2015; Kazdin, Mazurick, & Siegal, 1994; Thapar et al., 2015). Psychodynamic therapy is considered to be more appropriate when the presenting problem is emotional or internalizing as opposed to behavioural and externalizing (Fonagy, 2015); it may follow that when a youth presents with both types of issues, the emotional component could be addressed through individual treatment,

and the behavioural through a family-based approach. However, long term psychodynamic interventions are rarely covered by insurance or public healthcare options, and thus not a realistic option for many families.

5.7.3 Cognitive Behavioural Approaches

Cognitive Behavioural Therapy's effectiveness for this population and problem was not addressed directly, nor did the research conducted in the overview and implementation analysis address the theory that a combination of this type of treatment and medication will effective in the reduction of behavioral disorders in youth (Lochman et al., 2011; Thapar et al., 2015). One study contained in the implementation analysis did include a combined FFT and CBT group, but the outcomes were not significantly different from the FFT only group (Waldron, Slesnick, Brody, Turner, & Peterson, 2001). It is beyond the scope of the research conducted in this thesis to determine whether or not CBT will or will not be beneficial if it is combined with FFT.

5.7.4 Residential Treatment

Youths who have been sent to inpatient or custodial treatment settings are not often treated with FFT, and this thesis does not directly address this population because severe or persistent mental health symptoms will likely preclude a youth from participation in the intervention (Alexander, 2013). Children and youth behave and exist within an environment and context, and some version of family treatment may be recommended upon release or even whilst still in the facility undergoing treatment. To release an individual and return him or her to the family setting without appropriately addressing the issues that are likely to present in the home will often result in adverse or unsustainable outcomes and a greater chance for recommitment.

5.7.5 Juvenile Justice

Juvenile Justice encompasses a number of different approaches and treatment options, and FFT has been increasingly recommended for those returning to the home after being in secure custody or mandated placement such as a group home or juvenile incarceration facility.

Additionally, FFT is often required in conjunction with probation. Given that FFT is primarily concerned with the reduction of recidivism, it is indicated by many providers, courts and stakeholders as a logical choice in these circumstances. One of the benefits of FFT is the absence of possible peer effects or stigmatization that are often present in other juvenile justice interventions. Dishion, McCord, and Poulin (1999) pointed out that interventions meant to reduce and treat disruptive, deviant or behaviorally disordered behaviours in youth may have adverse effects. Their analysis shed light on the fact that any positive effects might be undercut by the peer influences present when youth are treated within their deviant peer groups. Continued research in this area strongly suggests that the presence of peer effects and deviancy training is an ongoing and very real problem, and that new approaches must be utilized and developed in order to avoid these occurrences (Dishion & Dodge, 2005). It is important to remember that FFT is not a juvenile justice intervention, rather it is viewed as a complementary treatment option to those already (or at imminent risk of) involvement in the juvenile justice system.

5.7.6 School Interventions

Current school intervention strategies and models do not appear in the evidence relating to FFT. No relationship or connection between FFT and school-based approaches is apparent or identified in the literature or trials of FFT. This is troubling in light of the fact that a child spends approximately eight hours a day in a school setting, more when involved with extracurricular activities such as sports, drama or clubs. In truth there may be a greater chance of a teacher or coach noticing troubling behaviours or symptoms than any other adult in his or her life. Additionally, teachers and other school staff may have more opportunities to observe how students interact with others or exhibits symptoms. A large proportion of acting out and behavioural disorders (i.e., fighting, bullying, opposition to authority, cheating, truancy, selling

or using of substances) manifest in the classroom or in a school setting. Teachers, coaches, counselors, club leaders and similar could be seen as valuable resources for identifying and modifying these behaviors. These adults and authority figures may be invaluable sources of support and help for a young person and should therefore be included in treatment. The family model advocated by FFT, which is based to a large degree on Minuchin (Minuchin, 1974) and Haley (Haley, 1976), allows for – and in some cases should demand – engagement with members of a child's environment and people (such as teachers and other school staff) outside the strict confines of immediate family members.

5.8 Implementation and FFT

FFT falls squarely into the category of systemic approaches, as it is built on a foundation of systems theory and includes both strategic and structural family approaches to treatment. As with many types of treatment, structural and strategic therapy has gone in and out of favour over the past few decades since its creation and development. FFT is but one of the treatment modalities aimed at treating the dysfunctional family that incorporates the theoretical signatures of Minuchin and Haley (Haley, 1976; Minuchin, 1974). It was designed while the influence of these modalities were gaining in strength and popularity. The approach that negative actions and disorders are likely a result of a dysfunctional or troubled family system is now a commonly accepted concept, but at the time was considered to be novel and even revolutionary. FFT was developed in this context, and has remained in many ways faithful to its theoretical base. While treating the family as a unit, and interactions as patterns of communication and relatedness, is a laudable and evidence supported approach, FFT may not be the most ideal means of doing so. What came to light through the overview and then was borne out by the implementation analysis is that it may not be an easily replicable or generalizable model. As with such other systemic approaches such as MST, the evidence in support of FFT continues to be inconclusive. Some

literature points out that family therapy (even strategic and structural therapy) does not have preordained and specified activities, but is rather a flexible framework for approaching the variable points and needs of each family (Haley, 1976; Lipsey, Chapman, & Landenberger, 2001). In fact, the designers of strategic and structural therapy *intend* for the approaches to be viewed more as a framework and support than a rigid model (Haley, 1976, Minuchin, 1974). FFT has been demonstrated throughout this thesis as attempting, in many ways ineffectually, to straddle the line between a manualized and pre-determined therapeutic intervention and a theory-based approach to families. There are detailed guidelines for providers, and supervision and oversight, but there is also an expectation to alter the methods and approach dependent upon the individual, family, and presenting problem. What the model lacks is any guidance around how to navigate between these two things.

There exists in implementation science a tension between fidelity - which is defined as the degree to which an intervention is delivered as it was specifically designed, intended, and perhaps manualized, and adaption, which refers to how an intervention is delivered or altered to fit different contexts in natural conditions in a real-world setting (Gottfredson & Gottfredson, 2002; Mihalic, 2004). Adaptations may be planned ahead of time, or adjustments may be made in the moment as an intervention is being delivered (Moore et al., 2013). Adaptations may be made due to ethnic, cultural, social or other contextual differences. These adaptations do not necessarily mean the intervention is no longer effective, but adaptions may also lead to the validity of the evidence base for the intervention base being called into question. There is no need to see model adaptation as oppositional to fidelity and adherence. One viewpoint suggests that adaptation is a threat or impediment to successful implementation (Mihalic, 2004). However it may be more useful and accurate in some cases to view adaptation, if and when the adaptions

and changes are made with a purpose, care, and an underlying fidelity to the theory and function of an intervention, as a needed component of successful delivery (Hawe, Shiell, & Riley, 2004). In the case of FFT, it may be that when there is a minimal recognition of contextual and participant differences, the model becomes less effective. In short, it may be that FFT requires increased context-specific adaptation, which need not be seen as a lack of fidelity or adherence.

5.9 Allegiance effects and bias throughout FFT research

A unifying element in both the overview and implementation analysis is the likely presence of significant allegiance bias in the trials and reviews of FFT. It is impossible to overlook the fact that a majority of the research has been conducted by those who have an interest, either financial or ideological, in the outcome of the trials, and in the manner in which those trials are reported and reviewed. The cynical view suggests that this allegiance bias is both intentional and nefarious, meaning that the developers are well aware of what they are doing, and how they are presenting evidence, and it is done in an attempt to increase uptake of their intervention. It is not entirely borne out by the evidence that this cynical view applies to FFT research, although the evidence does indicate that those reviews and trials conducted or authored by individuals with a stake in FFT had far more positive results on core outcome measures than those authored or conducted by independent researches.

The alternative *non*-cynical interpretation of the disparity in results is that developers or designers of the model are more familiar with its method and are as a result of this familiarity more likely to implement it with fidelity and adherence. The oversight and supervision, two elements which were shown through the implementation analysis to have bearing on core outcome measures, were more comprehensive when the model was delivered by designers or developers of FFT. This interpretation - the 'high-fidelity view' - would suggest that there is nothing inherently wrong with the study or results, but that those results would lack external

validity and not be generalizable or applicable to other settings. The efficacy trials of an intervention, when trials are performed under ideal circumstances, must be replicable in effectiveness trials, which examine and measure how an intervention performs in a real world setting, when the designers of the model are not providing strict adherence, support, and oversight.

In many ways, the issue is not *why* there exist such differences in outcomes, but what their existence implies for the research base as well as for the success of the model. Bias may very well be the cause of these differences, and whether the biases are conscious and intentional, or unconscious, is not truly the issue. The more pressing concern is whether the positive reports relating to FFT, both in trials and reviews, can be considered as reliable and trusted sources of information and evidence. This thesis has demonstrated that the evidence in support of FFT is not consistent nor does it achieve the level of certainty that should be demanded when implementing interventions relating to such a vulnerable population as at-risk youth and their families.

5.10 Potential Harms of FFT

Possible harms should not be limited to psychological harm or increased likelihood of recidivism. When considering the possibility of harm, it is useful to consider that the concept can be broken into five separate areas: direct harm refers to the possibility that an intervention may have directly harmful (usually health related) effects; psychological harms may be considered as direct or indirect and are defined as negative psychological impact; equity harms exist when an intervention creates worsening health or societal inequalities and notes that even when individuals are not worse off, inequality itself is harmful; group and social harms often refer to peer effects and deviancy training, and also may occur when a particular group is singled out or stigmatized; finally, opportunity costs refers to any benefits that may be lost as a result of

allotting resources to ineffective interventions (Lorenc & Oliver, 2014). It must be kept in mind that there are inherent harms in taking up an intervention that does not have a strong and provable evidence base. The possibility of harm and negative effects must be considered when implementing or taking up an intervention, and these possibilities must be thought of through a broad and nuanced lens as they may not be easily measured or identified through more traditional means.

Neither the overview and implementation analysis identified the presence of any harms relating to FFT; harms were not specifically identified, reported, or measured on in any review or study. This in and of itself does not rise to the level of determinative evidence, but it would be folly to assume that because no harms were reported, no harms exist. Many of the reviews were shown to be incomplete, potentially biased, or authored long before the advent of guidelines relating to reporting or methodology. The same is true of the trials included in the implementation analysis. What this suggests is that harms must be more broadly defined, and that going forward, reviewers and trialist alike must pay closer attention to these issues and report them more comprehensively when their presence is indicated.

5.11 Strengths and limitations of the overview

The overview of FFT was the first overview to be conducted of FFT and has a stronger and more rigorous design than many existing reviews. The overview is, unlike many of the reviews included in it for analysis, in line with Cochrane guidelines as discussed in section 2.6.1 of this thesis. Also, the Cochrane Handbook has within it a specific section relating to the completion of Cochrane Overviews of Reviews. To adopt the Cochrane guide in its entirety would have excluded the majority of the FFT literature and research, so it was used to inform and focus rather than determine absolutely. Thus, the overview is not defined as a Cochrane Overview of Reviews, nor is it published with Cochrane. There was a clear a priori protocol for

the overview and the methods used in searching, data extraction and synthesis were clearly detailed; potential for biases was reported and there existed a high level of transparency throughout. The design and research method for the overview, as well as inclusion and exclusion criteria, was presented in protocol form and approved by three unbiased academic advisors, the data extraction was double coded by Nadine Pfeiffer and Gregory Norfleet and any disagreements resolved by Paul Montgomery. Throughout the overview process, authors were contacted for clarification and for unpublished or grey literature. Overviews are considered helpful and appropriate in the presentation and collation of certain types of information, but they are by nature quite broad. Details relating to individual characteristics, such as ethnicity, age and socio-economic status may be lost. Furthermore, differences in implementation, reporting or process may also not be captured. Thus, it is important to keep in mind that much of the minutiae and variation that can be identified when reporting on trials or, one step out, in the reviews of said trials, may not be picked up through overview methodology (Ballard & Montgomery, 2017). Rather than being forced to sort through the 31 published reviews of FFT and attempting to evaluate and determine which reviews are the most reliable or salient, the overview can distil the existing research into one accessible and systematic document that simplifies and clarifies an intervention's evidence base (Baker et al., 2014).

5.12 Reflexivity

Reflexive writing and practice is designed for, and implemented when, practitioners and social scientists in particular attempt to understand and perhaps challenge dominant concepts, assumptions and over-arching beliefs surrounding the concept of objectivity (Bolton, 2018; Dean, 2017). Reflexivity may allow the writer to consider what can be changed, and how to work with what cannot, how to further value different and possibly contrary perspectives, how others may perceive you and how these perceptions are impactful, and finally how it may be

possible to counteract social, political and cultural structures that may be taken for granted (Bolton, 2018). It is important to raise these questions and concerns, especially in the social sciences, as positionality is vital to a more in depth and complex understanding of social problems. In social care, health and welfare settings, practitioners and researchers are urged by extant literature and scholarship to always be critical and to apply perspective to themselves and their practice, and to "explore the resources that brings to bear on scientific inquiry, including their social location, social positioning, and motivations" (Dean, 2017, p. 7).

It became increasingly evident as this research was being conducted that some reflection on the process through which I was undertaking this project would add a significant and necessary contextual element to the work. Thematically, issues relating to conflicts of interest and allegiance bias run throughout the thesis, and presenting these issues as they relate to other researchers whilst refraining from transparency around my own positionality would be disingenuous.

I arrived at my decision to investigate FFT from a clinical background, I had over a decade of direct clinical practice as a Licensed Clinical Social Worker (LCSW), in Louisiana, New York, Texas and Oregon, a designation that in the United States requires at least 2,000 hours of direct clinical practice and 100 hours of clinical supervision. I underwent instruction and clinical training at Smith College School for Social Work and have worked in the forensic and child welfare systems of New York, Connecticut, New Jersey, Texas and Louisiana, where my focus has always been on children and youth who were at risk or involved in the juvenile justice system. My work in Louisiana after Hurricane Katrina contributed in many ways to my transitioning from direct service to research and academia. I came to perceive systemic failures in agencies that were leading directly to the inadequate and ineffective treatment of children,

youth and families in need. It became clear that a greater understanding of the evidence and research contributing to stakeholder, practitioner and policymaker decisions regarding possible treatment options was needed. I chose to pursue my doctoral degree in an effort to achieve this for my own practice but also with the ultimate goal of being able to contribute to a more transparent, complete, and well-balanced research base in the area of behaviorally disordered youth and family treatment options.

The situational, contextual, and demographic elements present in post-Katrina Louisiana were at the time, and continue to be, completely different from the white, educated, middle class populations in which FFT was primarily trialed and delivered. Not only were there contextual differences, there was also a lack of diversity within our own FFT team, and I myself did not align either ethnically or culturally with the majority of my clients.

There has not been enough analysis of how the intervention may or may not work in specific settings to adequately determine when and where to adopt its use. When the agency I worked for in Louisiana chose to adopt FFT, it was done so without a complete picture of the evidence. Perhaps it does not reflect well on me as a practitioner that I adopted an intervention for treatment that had such a disparate research base. However, it must be noted that, no matter how well trained a practitioner may be, they are not always, or even often, trained in the particularities and idiosyncrasies of evidence-based practice. It was, in a sense and in fact, sold to the team in Louisiana as a model that we could use, that would be supported, that would work for our clients. In reality, we were not given consistent or intensive support, nor were we supported by FFT providers who knew anything about the region or population. We were faced with a lack of resources and a lack of infrastructure that most certainly caused detriment to our practice, and we were not given viable options for circumventing these problems.

Eventually it became clear that FFT, as it was being implemented by our team and in our specific context, was not an effective or appropriate means of addressing the problems of juvenile delinquency and behavioral disorders. The agency and clinicians within it were not creating significant positive change for the youth and families we were working with. I began to see the stark contrast between what FFT was designed and promoted as being able to accomplish and what it could, in the real world, actually achieve.

This should serve, in part, as a warning to understand the context of where and how an intervention is to be implemented fully before subscribing to its use. My own experience delivering the model and observing it in practice led to an increased desire to obtain a complete and honest interpretation of the data relating to FFT which has manifested as this thesis's exploration and analysis of the extant evidence supporting its use.

The completion of the overview was the first step towards a greater understanding of the evidence, and after the publication of the overview, the editor for the journal in which it appeared, *Research on Social Work Practice* (RSWP), Dr. Bruce Thyer, was contacted by Dr. Mike Robbins, one of the original developers and designers of the FFT model. Dr. Robbins requested that the article be pulled from publication. In his initial email, he states that "from [his] perspective, the conclusions that are being drawn are irresponsible and misleading. Moreover, these inaccurate conclusions have the potential to cause harm for thousands of families that are currently receiving FFT in real world settings" (Communication from Michael Robbins). He goes on to suggest an immediate retraction of the article, a suggestion which the editor summarily rejected on the grounds that the journal is highly regarded and peer reviewed; the editor responded to Dr. Robbins' request by pointing out that doing so would "obviate the blind peer review process" (Communication from Bruce Thyer). Dr. Robbins was invited to author a

response to the overview, which I could then address. The nature of my interactions with Dr. Robbins, and the tone of his article, was aggressive and dismissive. I felt on some occasions to be under personal attack, and the experience was distressing on a number of levels. Robbins then authored a critique and call for revision of the article, which asserted that the conclusions put forth in the overview were inaccurate and unsound scientifically (Robbins & Turner, 2018). Rather than causing me to doubt my initial methods, research, and conclusions, the aggressive and strident nature of both his correspondence and article only served to lend credence to my finding that there may indeed be a strong allegiance bias present in FFT research and literature. Robbins disclosed a significant conflict of interest, as he is employed by FFT LLC which is the organization tasked with the dissemination of FFT into community and practice settings. His co-author, Dr. Charles Turner, has been involved with FFT research for over 45 years. In their conflict of interest statement, they are careful to state that Dr. Turner receives no financial compensation for his work and no financial affiliation with FFT LLC. Dr. Turner's career is based almost entirely upon the success and uptake of the model and presenting himself as not having a conflict of interest is a deliberate attempt to obfuscate this fact.

The rebuttal to Robbins and Turner (2018) (Weisman & Montgomery, 2019) states the following: "Robbins and Turner discuss the reviews conducted by FFT developers and independent authors. Across these groups we identify allegiance effects" (p. 359). Conflict of interest occurs when there is even one member of a team who holds an interest in the intervention, which is not limited to direct financial gain but can include future career progression, reputation, publications, and "may be linked to copywrites, royalties, research funding and income generated from the distribution of programs" (Eisner, 2009, p. 170). Positive effects of an intervention during a study conducted by 'developers-as evaluators' may be

attributed to either a) the possibility that the quality of implementation is of a higher standard when it is delivered by a developer of the intervention or b) disparity in outcomes between studies delivered by developers versus independent researchers that can be attributed to systemic bias relating to an ongoing conflict of interest by a champion of the program (Eisner, 2009). Having a vested interest in the outcome and dissemination of FFT is, we believe, a concern with these data. Referring only to financial conflicts of interest understates the issues involved. Allegiances and loyalty exist on a number of different levels and manifest themselves in a number of ways" (Weisman & Montgomery, 2019, p. 359).

The result of this process was to further my conviction that continued research into FFT was necessary, as well as my continued and heightened awareness of any unconscious bias I may have *against* the intervention. Ideological biases may be present for any number of reasons, leaning in either direction, and my interactions with the designers of FFT could not be ignored as a possible source of unintended and unconscious bias. In order to ensure this was not the case, I ensured that there existed a heightened level of transparency around every step of the research process. I then went on to complete the implementation analysis which, in conjunction with my own past experiences as a practitioner of the model in New Orleans, served to reconfirm that implementation and delivery of the model may be at the core of whether it is or is not effective, and for whom this is the case. I did not, at any point in my research, disparage or devalue FFT. Rather I attempted to present a complete and unbiased picture of the evidence. The implementation analysis shed further light on possible areas of concern and investigation of FFT, and I began to consider how best to continue my post-doctoral research into FFT in a post-doctoral capacity. I hoped to further develop my research through qualitative methodology, not in any attempt to disprove the model, rather to understand it more fully and contribute to the

research base in a productive and honest manner. I began contacting practitioners in an effort to foster positive relationships that would possibly assist me in conducting a qualitative study examining the decisions and methods they employed in the delivery of the model. However, once this process began, Dr. Robbins once again made a sustained effort to curtail my research (Weisman, 2018; Weisman & Montgomery, 2019).

My contact with practitioners was accompanied by a covering letter which outlined the research completed prior and explicated the purpose and scope of the proposed project. This letter states that it would be "beneficial to explore the methods, rationale and experiences of the therapists and clinicians who are working with families to determine how they make changes or adapt the model with a 'difficult to reach' and severely underserved population…The therapists who are tasked with this work will provide a wealth of information regarding to how the model truly functions when clinicians have to use their own judgements and personal experiences to inform their decision making and treatment of any given case" (Communication to FFT sites from Paul Montgomery & Clio Weisman, February 2019). Initially my efforts were met by those I spoke to with openness and the potential to participate in research. However, Dr. Robbins then requested that I cease all contact with these agencies, and sent the following email to my supervisor, Dr. Paul Montgomery: "I represent FFT LLC, and in recent weeks I have received several calls and emails about a project that you and Clio Weisman are proposing to Functional Family Therapy sites in Louisiana…FFT LLC supports and certifies these teams in Louisiana. We have contracts with them that state any research at the site must be done in collaboration between FFT LLC and the local agency. So that you are also aware, our clinical tracking system, the CSS, may only be used for clinical and quality assurance purposes, unless otherwise approved by FFT LLC (Communication from Michael Robbins, 05 February, 2019). Dr.

Montgomery and myself, after discussion, determined that the most politic and useful means of approaching the issue was to have a conversation with Dr. Robbins in which we could better explain the project, and hopefully de-escalate any perceived animosity. We emailed him directly explaining the goals of the proposed project and arranged a Skype conversation.

During this conversation, Dr. Robbins stated repeatedly, and defensively that he had no involvement with many FFT studies, and that the research being conducted was independent. However, his prior emails and communication as well as what he stated explicitly in the Skype meeting indicated that he and FFT LLC had to have at least peripheral involvement in all FFT research, that they in fact had to give permission for any research to be conducted. This is absolutely *not* independent evaluation or research and is therefore far more likely to be subject to bias. Dr. Robbins did not waver in his position despite our making it very clear that were not wishing to access client records or protected documentation, that this project would focus on the experiences of the therapists (who it should be noted are not employed by FFT LLC, rather by their home agencies). He agreed, finally, to contact the advisory board of FFT and discuss the conversation and proposed project with them. We then received the following communication from Dr. Robbins:

> As I promised on our call last week, I reached out to sites that FFT LLC contracts with in Louisiana and to our FFT LLC advisory group. I also shared details of our conversation, including how this project fits within a program of studies that you have been conducting on the FFT model as well as your perspective about evidence-based models and future directions for the field of psychotherapy/family therapy.
>
> We do not have any sites that were interested in participating in this evaluation. The members of our FFT advisory group also declined to participate. We respectfully request that you refrain from further contact with therapists and agencies in Louisiana. As I noted in our conversation, we have existing contracts with sites that require the agreement between the site and FFT LLC about any potential research projects and it would be putting therapists in a potentially risky and awkward situation with their home agency if they are approached directly. (Communication from Michael Robbins, 01 March, 2019)

Dr. Robbins followed this email up with the statement that "all of our contracts… require mutual agreement of the site and FFT LLC for research to occur… we will not be giving our agreement to sites for this research. Thus, we ask that you and your student stop contacting FFT LLC sites for your project. Otherwise, we will be pressed to take to appropriate legal steps" (Communication from Michael Robbins, 04 March 2019). The level of Dr. Robbins' animosity and involvement with FFT research is clearly present in these communications, as well as the fact that this does in truth violate the nature of independent research, a point which he was not amenable to. With these actions, FFT LLC effectively halted my *truly* independent research into the model, as they were in effect attempting to prevent any research that they did not expressly condone; their threat of legal action, while being baseless and surmountable, amounted to an attack and to further evidence of bias.

Standing in the way of independent research is on its face problematic, and I believe Dr. Robbins and FFT LLC to be entirely unethical in their doing so. His inaccurate representations of my position, goals and even theoretical base led to him acting in manner that further exemplified a lack of impartiality. It was impossible not to interpret his behaviour and conduct as aggressive and closed to any perspective that conflicted with his own. Reflexivity demands that we be cognizant of our own position and relationship to our research, and my sense of frustration and even anger buoyed my determination to prove the validity of my conclusions regarding FFT.

With the benefit of hindsight, it may have been beneficial to contact Dr. Robbins before any research into FFT was undertaken. His cooperation and support certainly would have allowed access to records and resources that may have been helpful, but in order to achieve that, I am certain compromises would have been made. After the publication of the overview, and his vitriolic response to it, I found myself disinclined to give him the benefit of the doubt, or to

respect his position, as I viewed him as in direct opposition to my intentions, and I had to remain very careful to keep my personal experiences from colouring the work. There may have been a benefit to making more of an effort to engage Dr. Robbins, but I still believe that his position is not one of neutrality, and it fundamentally goes against my conviction that independent research is both valuable and necessary. Through his actions, Dr. Robbins served only to contribute to the mounting evidence that supported my belief – based on empirical data and now the actions of FFT LLC – that allegiance bias, conflicts of interest, and perhaps even ethical misconduct, were and continue to be substantial elements present in current FFT research.

5.13 Conclusions and suggestions for further research

The research conducted within this thesis presents an evidentiary picture of FFT that does not meet a high standard of completeness, impartiality, or depth. Through the overview and implementation analysis, issues have been identified that are significant causes for concern. Chief among these concerns are the presence of bias and the model's apparent lack of generalizability across different contexts.

This thesis indicates the need for truly independent investigations into FFT. What is shown in the overview and implementation analysis is a clear relationship between positive outcomes and the involvement of designers or others invested in FFT being taken up, disseminated and presented as an effective treatment option. High quality, scientifically sound and impartial research would be greatly beneficial in strengthening the evidence relating to FFT. Additionally, contextual elements and differences need to be further examined in order to evaluate where, and for whom, the intervention is appropriate, effective and beneficial. The field of social work must continue to examine and evaluate best practices and treatment options, and the evidence portrayed throughout this thesis contributes to this endeavor. Also, the methodology and rigour employed herein may be extrapolated for use in other contexts and to address

concerns relating to other interventions or therapeutic models. Future research should also focus on untangling the presence of allegiance bias, and determining whether the extant research reports outcomes that are inextricably linked to these biases. The lessons of this thesis can reach beyond FFT and be used to push the field of evidence supported intervention towards increased levels of transparency and continued efforts to reduce bias in reporting, reviewing, and authorship.

At this time, there is little evidence of whether FFT may be appropriate for lower income and culturally diverse populations, yet these are precisely the contexts in which it is being increasingly delivered. To do so without a strong and consistent evidence base supporting its use is irresponsible and may be eventually demonstrated as harmful. Also, considering that behavioural problems exist across all socioeconomic and cultural groups, including lower or middle income countries, it would be greatly beneficial to investigate whether FFT, or some adaptation of it, might be useful in addressing these problems in less developed areas.

Given the issues identified through both the overview and implementation analysis, it is vital that future research of FFT be conducted by those who have no allegiance, in any form, to the model. This independence must be supported, and if FFT LLC has, as they state, an absolute confidence in their model, it seems they might be willing and eager to support further research without *any* involvement on their part.

Further, the adequate examination of FFT requires in-depth and qualitative studies, once again conducted not by those who may have a vested interest, to broaden and open up the research base, specifically around implementation concerns and the delivery of the model across varied contexts and settings. Racial and ethnic composition of participants and therapists should be examined in greater detail. When considering the damaging impact juvenile behavioural

problems have on multiple levels, it is necessary to implement interventions that are supported by unbiased and rigorous evidence. FFT does not, at this time, meet this standard, and this is borne out by the evidence presented throughout this thesis.

6 REFERENCES

Achenbach T. M. (1991). *Manual for the Child Behavior Checklist/4–18 and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment*. Hillsdale, NJ: Erlbaum.

Alexander, J. (2004). Blueprints for Violence Prevention: Functional Family Therapy, Denver, Colorado. Retrieved from https://www.ncjrs.gov/pdffiles1/ojjdp/204274.pdf

Alexander, J. F., Waldron, H. B., Robbins, M. S. & Neeb, A. A. (2013). *Functional Family Therapy for Adolescent Behavior Problems*. Washington, DC: American Association.

Alexander, J. F. & Parsons, B. V. (1973). Short term behavioral intervention with delinquent families: Impact on family process and recidivism. *Journal of Abnormal Psychology, 81*, 219- 25.

Alexander, J. F., Barton, C. Schiavo, R. S. & Parsons, B. V. (1976). Behavioral intervention with families of delinquents: Therapist characteristics and outcome. *Journal of Consulting and Clinical Psychology, 44*, 656-64. doi: 10.1037/0022-006X.44.4.656

Alexander, J. F., & Sexton, T. L. (2002). Functional family therapy: A model for treating high-risk, acting-out youth. In F. W. Kaslow (Ed.), *Comprehensive handbook of psychotherapy: Integrative/eclectic, Vol. 4* (pp. 111-132). Hoboken, NJ: John Wiley & Sons Inc.

Alexander, J., Pugh, C., Parsons, B., & Sexton, T. (2002). *Functional Family Therapy*. Boulder, CO: Center for the Study and Prevention of Violence, Institute of Behavioral Science, University of Colorado Boulder.

Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). Benefits and costs of prevention and early intervention programs for youth. Olympia: Washington State Institute for Public Policy.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text Revision). Washington, DC.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC: Author.

Antman, E. M., Lau J., Kupelnick, B., Mosteller, F., & Chalmers T.C. (1992) A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: Treatments for myocardial infarction. *JAMA, 268*, 240-48.

Armelius, B. & Andreassen, T. (2007). Cognitive-behavioral treatment for antisocial behavior in youth in residential treatment. *Cochrane Database of Systematic Reviews* (Online). 4. CD005650. 10.1002/14651858.CD005650.pub2.

Arthur, M. W. & Blitz, C. (2000). Bridging the gap between science and practice in drug abuse prevention through needs assessment and strategic community planning. *Journal of Community Psychology, 28*, 241-55. doi: org/10.1002/(SICI)1520-6629(200005)28:3<241::AID-JCOP2>3.0.CO;2-X

Austin, A. M., Macgowan, M. J., & Wagner, E. F. (2005). Effective family-based interventions for adolescents with substance use problems: A systematic review. *Research on Social Work Practice, 15*, 67-83. doi:10.1177/1049731504271606

Baglivio, M. T., Jackowski, K., Greenwald, M. A., & Wolff, K. T. (2014). Comparison of multisystemic therapy and functional family therapy effectiveness: A multiyear statewide

propensity score matching analysis of juvenile offenders. *Criminal Justice and Behavior*, *41*, 1033-1056. doi:10.1177/0093854814543272

Barton, C., Alexander, J. F., Waldron, H., Turner, C. W. & Warburton, J. (1985). Generalizing Treatment Effects of Functional Family-Therapy - 3 Replications. *American Journal of Family Therapy, 13*, 16-26. doi: 10.1080/01926188508251260

Baker, P. R. A., Costello, J. T., Dobbins, M., & Waters, E. B. (2014). The benefits and challenges of conducting an overview of systematic reviews in public health: a focus on physical activity, *Journal of Public Health*, *36*, 517- 21. doi.org/10.1093/pubmed/fdu050

Baldwin, S. A., Christian, S., Berkeljon, A. & Shadish, W. R. (2012), The effects of family therapies for adolescent delinquency and substance abuse: A meta-analysis. *Journal of Marital and Family Therapy*, 38: 281–304. doi:10.1111/j.1752-0606.2011.00248.x

Ballard, M., & Montgomery, P. (2017). Risk of bias in overviews of reviews: a scoping review of methodological guidance and four-item checklist. *Research Synthesis Methods*, *8*, 92–108. doi: 10.1002/jrsm.1229.

Bandura, A., & Walters, R. H. (1959). *Adolescent Aggression*. New York: Ronald Press.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.doi.org/10.1037/0033-295X.84.2.191

Barnett, J. E. (2011). Utilizing technological innovations to enhance psychotherapy supervision, training, and outcomes. *Psychotherapy, 48*, 103–08.

Barnoski, R. (2004, March) *Assessing risk for re-offense: Validating the Washington State Juvenile Court Assessment*. Retrieved from http://www.wsipp.wa.gov/ReportFile/868/Wsipp

Barton, C., Alexander, J. F., Waldron, H., Turner, C. W. & Warburton, J. (1985). Generalizing Treatment Effects of Functional Family-Therapy - 3 Replications. *American Journal of Family Therapy, 13*, 16-26. doi: 10.1080/01926188508251260

Beck, A. T. (1967). *Depression: Causes and treatment*. Philadelphia, PA: University of Pennsylvania Press.

Becker, L. A., & Oxman, A. D. (2008). Overviews of reviews. In Higgins, J. P. T., & Green, S. (Eds.), Cochrane handbook for systematic reviews of interventions: Cochrane book series. Chichester, England: John Wiley & Sons.

Beech, A. R., Bourgon, G., Hanson, K., Harris, A. J., Langton, C., Marques, J., Miner, M., Murphy, W., Quinsey, V., Seto, M., Thornton, D., & Yates, P.M. (2007). *Sex Offender Treatment Outcome Research: CODC Guidelines for Evaluation Part 1: Introduction and Overview*. Ottawa, ON: Public Safety Canada.

Behan, J., & Carr, A. (2000). Oppositional defiant disorder. In A. Carr (Ed.), *What works for children and adolescents? A critical review of psychological, interventions with children, adolescents and their families.* (pp 102-130). London: Routledge Browne.

Bender, K., Tripodi, S. J., Sarteschi, C. & Vaughn, M.G. (2011). A meta-analysis of interventions to reduce cannabis use. *Research on Social Work Practice*, *21*, 153-64. doi: org/10.1177/1049731510380226

Bolton, G. (2018). *Reflexive practice: Writing and professional development*. London, UK: Sage.

Bonell, C., Oakley, A., Hargreaves, J., Strange, V. & Rees, R. (2006). Assessment of generalisability in trials of health interventions: suggested framework and systematic review. *BMJ, 333*, 346-49. doi.org/10.1136/bmj.333.7563.346

Bonell, C., Fletcher, A., Morton, M., Lorenc, T., & Moore, L. (2012). Realist randomised controlled trials: a new approach to evaluating complex public health interventions. *Social Science and Medicine, 75*, 2299-306. doi: 10.1016/j.socscimed.2012.08.032.

Bonell, C., Jamal, F., Melendez-Torres, G. J. & Cummins, S. (2015). Dark logic: Theorising the harmful consequences of public health interventions, *Journal of Epidemiology and Community Health, 69*, 95.

Bongers, I. L., Koot, H. M., van der Ende, J. & Verhulst, F. C. (2003). The normative development of child and adolescent problem behavior. *Journal of Abnormal Psychology*, *112*, 179-92. doi: 10.1037/0021-843X.112.2.179

Bornstein, M. H., Hahn, C. S., & Haynes, O. M. (2010). Social competence, externalizing, and internalizing behavioral adjustment from early childhood through early adolescence: developmental cascades, *Development and Psychopathology, 22*, 717-35. doi: 10.1017/S0954579410000416.

Bowlby, J. (1969). *Attachment and Loss*. New York; Basic Books.

Breunlin, D. C., Breunlin, C., Kearns, D. L., & Russell, W. P. (1988). A review of the literature on family therapy with adolescents 1979 - 1987. *Journal of Adolescence, 11*, 309-334. doi:10.1016/S0140-1971(88)80032-0

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA; Harvard University Press.

Cabral, R. R. & Smith, T. B. (2011). Racial/ethnic matching of clients and therapists in mental health services: A meta-analytic review of preferences, perceptions, and outcomes. *Journal of Counseling Psychology,* 58, 537-554. doi: 10.1037/a0025266.

Caird, J., Sutcliffe, K., Kwan, I., Dickson, K., & Thomas, J (2015). Mediating policy-relevant evidence at speed: are systematic reviews of systematic reviews a useful approach? *Evidence & Policy, 11*, 81-97. doi.org/10.1332/174426514X13988609036850

Campbell, S. B., Shaw, D. S., & Gilliom, M. (2000). Early externalizing behavior problems: Toddlers and preschoolers at risk for later maladjustment. *Development and Psychopathology, 12*, 467-88. doi: org/10.1017/S0954579400003114

Carr, A. (2000). Evidence-based practice in family therapy and systematic consultation: Child-focused problems. *Journal of Family Therapy, 22*, 29-60. doi:10.1111/1467-6427.00136

CASP – Critical Appraisal Skills Programme. (n.d.). Retrieved from https://casp-uk.net/

Celinska, K., Furrer, S. & Cheng, C. C. (2013). An outcome-based evaluation of functional family therapy for youth with behavioral problems. *OJJDP Journal of Juvenile Justice*, *2*, 23-36.

Chamberlain, P. R. J. G. (1995). The effectiveness of family therapy in the treatment of adolescents with conduct disorders and delinquency. *Journal of Marital and Family Therapy, 21*, 441-459.

CJCA. The Council of Juvenile Correctional Administrators. (2018). Retrieved from http://cjca.net/Cochrane Comparing Multiple Interventions Methods Group. Retrieved from https://methods.cochrane.org/cmi/

Colman, I., Murray, J., Abbott, R. A., Maughan, B., Kuh, D., Croudace, T. J., & Jones, P. B. (2009). Outcomes of conduct problems in adolescence: 40 year follow-up of national cohort. *BMJ (Clinical research ed.)*, *338*, a2981. doi:10.1136/bmj.a2981

Costello, E J.., Erkanli, A., & Angold, A. (2006). Is there an epidemic of child or adolescent depression? *Journal of Child Psychology and Psychiatry*, *47*(12), 1263-1271.

Cottrell, D., & Boston, P. (2002). The effectiveness of systemic family therapy for children and adolescents. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 43*, 573-586.

Cunningham, P. B. et al. (2018). Midtreatment problems implementing evidence-based interventions in community settings. *Family Process, 58*, 287-04. doi.org/10.1111/famp.12380

Darnell, A. J. & Schuler, M. S. (2015). Quasi-experimental study of Functional Family Therapy effectiveness for juvenile justice aftercare in a racially and ethnically diverse community sample. *Child Youth Services Review, 50*, 75-82. doi: 10.1016/j.childyouth.2015.01.013

Dean, J. (2017). *Doing reflexivity*. Chicago, IL: University of Chicago Press.

Diamond, G. S., Serrano, A. C., Dickey, M. I. T. C., & Sonis, W. A. (1996). Current Status of Family-Based Outcome and Process Research. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*, 6-16.

Dishion, T., McCord, J. & Poulin, F. (1999). When Interventions Harm: Peer Groups and Problem Behavior. *The American Psychologist, 54*, 755-64. doi: 10.1037/0003-066X.54.9.755.

Dumas, J. E. (1989). Treating antisocial behavior in children: Child and family approaches. *Clinical Psychology Review, 9*, 197-222

Durlak, J. A. (1995). *Developmental clinical psychology and psychiatry, Vol. 34. School-based prevention programs for children and adolescents.* Thousand Oaks, CA; Sage Publications, Inc.

Durlak, J. A. (1998). Why program implementation is important. *Journal of Prevention and Intervention in the Community, 17*, 5-18.

Durlak, J. A. & DuPre, E. P. (2008). Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Clinical Psychology, 41,* 327-50, doi: org/10.1007/s10464-008-9165-0

Eisner, M. P. (2009). No effects in independent prevention trials: Can we reject the cynical view? *Journal of Experimental Criminology, 5*, 163-83, doi: 10.1007/s11292-009-9071-y

Elliott, D.S. & Mihalic, S. (2004). Issues in disseminating and replicating effective prevention programs. *Prevention Science, 5*, 47-53. doi.org/10.1023/B:PREV.0000013981.28071.52

Erikson, E. H. (1956) The problem of ego identity. *Journal of the American Psychoanalytic Association, 4*, 56-121. doi: 00.1177/000306515600400104

Fairburn, C. G. & Cooper, Z. (2011). Therapist competence, therapy quality, and therapist training. *Behaviour Research and Therapy, 49*, 373-78, doi: 10.1016/j.brat.2011.03.005

Farrington, D. P. W. B. C. (2003). Family-based prevention of offending: A meta-analysis. *Australian and New Zealand Journal of Criminology, 36*, 127-151.

Fergusson, D. M., Woodward, L. J., & Horwood, L. J. (1999). Childhood peer relationship problems and young people's involvement with deviant peers in adolescence, *Journal of Abnormal Child Psychology, 27,* 357-69. doi.org/10.1023/A:1021923917494

Fergusson, D. M., Horwood, L. J., & Ridder, E. M. (2005). Show me the child at seven: the consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology and Psychiatry, 46*, 837-49. doi: 10.1111/j.1469-7610.2004.00387.x

Filges T, Andersen D, Jørgensen, A. M. K. (2018). Functional Family Therapy (FFT) for Young People in Treatment for Non-Opioid Drug Use: A Systematic Review. *Research on Social Work Practice*, 28, 131-145. doi:org/10.1177/1049731516629802

Flicker, S. M., Waldron H. B., Turner, C. W., Brody, J. L. & Hops, H. (2008). Ethnic matching and treatment outcome with Hispanic and Anglo substance-abusing adolescents in family therapy. *Journal of Family Psychology, 22*, 439–47.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2011*) Implementation Research: A synthesis of the literature*. National Implementation Research Network, University of South Florida, Tampa, FL.

Fonagy, P. (2015). The effectiveness of psychodynamic psychotherapies: An update. *World Psychiatry, 14*, 137-50. doi: 10. 1002/wps.20235

Fonagy, P. & Target, M. (1994). The efficacy of psychoanalysis for children with disruptive disorders. *Journal of the American Academy of Child Adolescent Psychiatry, 33*, 45-55. doi: 10.1097/00004583-199401000-00007

Fonagy, P., Butler, S., Cottrell, D., Scott, S., Pilling, S., Eisler, I., et al. (2018). Multisystemic therapy versus management as usual in the treatment of adolescent antisocial behaviour (START): a pragmatic, randomised controlled, superiority trial. *The Lancet, 5*, 119-33. doi.org/10.1016/S2215-0366(18)30001-4

Freud, S. (1955). *The interpretation of dreams*. Philadelphia, PA; Basic Books.

Friedman, A. S. (1989). Family therapy vs. parent groups: Effects on adolescent drug abusers. *American Journal of Family Therapy, 17*, 335-47.

Garbarino, J. (1990). The human ecology of early risk. In S. J. Meisels & J. P. Shonkoff (Eds.), *Handbook of early childhood intervention* (pp. 78-96). New York, NY, US: Cambridge University Press.

Ghate, D. (2015). *The Family Links Nurturing Programme – Findings and recommendations from a strategic review and development project. Extended summary 2: Reviewing the Theory of Change for the Nurturing Programme*. Oxford: Family Links.

Ghate, D. (2018). Developing theories of change for social programmes: co-producing evidence-supported quality improvement. *Palgrave Communications, 4.* doi:10.1057/s41599-018-0139-z

Gibbs, L. & Gambrill, E. (2002). Evidence-Based Practice: Counterarguments to Objections. *Research on Social Work Practice, 12*(3), 452-476. https://doi.org/10.1177/1049731502012003007

Goldstein, A. P., Glick, B., Gibbs, J. C. (1998). *Aggression Replacement Training: A comprehensive intervention for aggressive youth.* Champaign, IL, US: Research Press

Goldstein, S. & Brooks, R. B. (Eds.). (2005). *Handbook of resilience in children*. New York; Kluwer Academic/Plenum Publishers.

Gordon, D. A. (1995). Functional Family Therapy for delinquents. In R. R. Ross, D. H. Antonowicz, & G. K. Dhaliwal (Eds.), *Going straight: Effective delinquency prevention and offender rehabilitation* (pp.163-178). Ottawa, Ontario, Canada: Air Training and Publications.

Gottfredson, D. C., Kearly, B., Thornberry, T. P., Slothower, M., Devlin, D. & Fader, J. J. (2018). *Scaling-up evidence-based programs using a public funding stream: A*

randomized trial of Functional Family Therapy for court involved youth. Paper submitted for publication.

Gottfredson, D. C. & Gottfredson, G. D. (2002). Quality of school-based prevention programs: results from a national survey, *Journal of Research in Crime and Delinquency, 39*, 3-35. doi: 10.1177/002242780203900101

Greenhalgh, T., Thorne, S., & Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews?. *European Journal of Clinical Investigation*, *48*, doi:10.1111/eci.12931

Gupta, A. K., Mongia, M., & Garg, A. K. (2017). A descriptive study of behavioral problems in schoolgoing children. *Industrial psychiatry journal*, *26*, 91–94. doi:10.4103/ipj.ipj_39_17

Haley, J. (1976). *Problem solving therapy*. San Francisco, CA; Jossey-Bass.

Hartling, L., Vandermeer, B. & Fernandes, R. M. (2014). Systematic reviews, overviews of reviews and comparative effectiveness reviews: a discussion of approaches to knowledge synthesis. *Evidence Based Child Health, 9,* 486-94. doi:10.1002/ebch.1968

Hartnett, D., Carr, A., Hamilton, E. & O'Reilly, G. (2017). The Effectiveness of Functional Family Therapy for Adolescent Behavioral and Substance Misuse Problems: A Meta-Analysis. *Family Process,* 56, 607–19. doi:10.1111/famp.12256

Hawe, P., Shiell, A. & Riley, T. (2004). Complex interventions: How "out of control" can a randomised controlled trial be? *BMJ, 328*, 1561-63. doi:10.1136/bmj.328.7455.1561

Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (2009). *Multisystemic therapy for antisocial behavior in children and adolescents*. New York: Guilford Press.

Henggeler, S. W., & Sheidow, A. J. (2003). Conduct disorder and delinquency. *Journal of*

Marital and Family Therapy, *29*, 505-522. doi:10.1111/j.1752-0606.2003.tb01692.x

Henggeler, S. W., & Sheidow, A. J. (2012). Empirically supported family-based treatments for conduct disorder and delinquency in adolescents. *Journal of Marital and Family Therapy*, *38*, 30-58. doi:10.1111/j.1752-0606.2011.00244.x

Higgins, J. P. T. & Green, S. (Eds.). *Cochrane Handbook for Systematic Reviews of Interventions.* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org.

Hoffman, T. C. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *The BMJ,* 348, g1687. doi: 10.1136/bmj.g168

Hogue, A., & Liddle, H. A. (2009). Family-based treatment for adolescent substance abuse: controlled trials and new horizons in services research. *Journal of Family Therapy, 31*, 126-154. doi:10.1111/j.1467-6427.2009.00459.x

Hops, H., Ozechowski, T.J., Waldon, H. B., Davis, B., Turner, C.J., Brody, J., & Barrera, M. (2011) Adolescent health-risk sexual behaviours: Effects of a drug abuse interventions. AIDS and Behaviour, 15, 1664-76.

Huey, S. J. J., & Henggeler, S. W. (2001). Effective community-based interventions for antisocial and delinquent adolescents. In J. N. Hughes, A. M. La Greca & J. C. Conoley (Eds.), *Handbook of psychological services for children and adolescents* (pp. 301-322). New York, NY: Oxford University Press.

Humayun, S., Herlitz, L., Chesnokov, M., Doolan, M., Landau, S. & Scott, S. (2017). Randomized controlled trial of Functional Family Therapy for offending and antisocial

behavior in UK youth. *Journal of Child Psychology and Psychiatry, 58*, 1023-32. doi:10.1111/jcpp.12743

James, C., Stams, G. J., Asscher, J. J., De Roo, A. K., & der Laan, P. H. (2013). Aftercare programs for reducing recidivism among juvenile and young adult offenders: A meta-analytic review. *Clinical Psychology Review, 33*, 263-274. doi:10.1016/j.cpr.2012.10.013.

James, S., Alemi, Q., & Zepeda, V. (2013). Effectiveness and implementation of evidence-based practices in residential care settings. *Children and Youth Services Review, 35*, 642-656. doi:org/10.1016/j.childyouth.2013.01.007

Johnson, T. J. & Friborg, O. (2015). The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: A meta-analysis. *Psychological Bulletin, 141*, 747-68. doi: 10.1037/bul0000015

Kazdin, A. E., Mazurick, J. L. & Siegal, T. C. (1994). Treatment outcome among children with externalizing disorder who terminate prematurely versus those who complete psychotherapy. *Journal of the American Academy of Child & Adolescent Psychiatry*, 33, 549-57.

Kempes, Maaike & Matthys, Walter & De Vries, Han & van Engeland, Herman. (2005). Reactive and proactive aggression in children: A review of theory, findings and the relevance for child and adolescent psychiatry. *European Child & Adolescent Psychiatry*. 14. 11-9. 10.1007/s00787-005-0432-4.

Klein, N. C., Alexander, J. F., & Parsons, B. V. (1977). Impact of family systems intervention on recidivism and sibling delinquency: A model of primary prevention and program evaluation. *Journal of Consulting and Clinical Psychology*, 45, 469-474.

Landenberger, N. A. & Lipsey, M. W. (2005). The positive effects of cognitive–behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology, 1*, 451-76. doi.org/10.1007/s11292-005-3541-7

Lipsey, M. W., Chapman, G. L., & Landenberger, N. A. (2001). Cognitive-behavioral programs for offenders. *The Annals of the American Academy of Political and Social Science*, *578*, 144-157. doi:10.1177/000271620157800109

Lipsey, M. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims & Offenders, 4,* 124-47. doi: 10.1080/15564880802612573.

Littell, J. H., Popa, M., & Forsythe, B. (2005). Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10-17 (Cochrane Review). *Cochrane Database of Systematic Reviews*, Issue 4, 2005.

Littell, J. H. & Shlonsky, A. (2009) Towards evidence informed policy and practice in child welfare. *Research on Social Work Practice, 70*, 723-25. doi: 10.1177/1049731509347886

Little, G. L., & Robinson, K. D. (1988). Moral reconation therapy: A systematic step-by-step treatment system for treatment resistant clients. *Psychological Reports*, *62,* 135-151.

Lochman, J. E., Powell, N. P., Boxmeyer, C. L., & Jimenez-Camargo, L. (2011). Cognitive-behavioral therapy for externalizing disorders in children and adolescents. *Child and Adolescent Psychiatric Clinics of North America, 20*, 305-20. doi: 10.1016/j.chc.2011.01.005.

Lorenc, T. & Oliver, K. (2014). Adverse effects of public health interventions: A conceptual framework. *Journal of Epidemiology and Community Health, 68*, 288-90. doi: 10.1136/jech-2013-203118

Loy, J. H., Merry, S. N., Hetrick, S. E. & Stasiak, K. (2012). Atypical antipsychotics for disruptive behaviour disorders in children and youths. Cochrane Database of Systematic Reviews. Version published: 12 September 2012. https://doi.org/10.1002/14651858.CD008559.pub2

Macintyre, S. & Petticrew, M. (2000). Good intentions and received wisdom are not enough. *Journal of Epidemiology and Community Health, 54*, 802-03. dx.doi.org/10.1136/jech.54.11.802

Mas, C. H., Alexander, J. F. & Barton, C. (1985). Modes of expression in family therapy: a process study of roles and gender. *Journal of Marital and Family Therapy, 11*, 411-15.

McCord, J. (1978). A thirty year follow up of treatment effects. *American Psychologist, 33*, 284-89.

McCulloch, A., Wiggins, R. D., Joshi, H. E., & Sachdev, D. (2000). Internalising and externalising children's behaviour problems in Britain and the US: relationships to family resources. *Children and Society, 14*, 368-83. doi.org/10.1111/j.1099-0860.2000.tb00191.x

McPherson, L. & Macnamara, N. (2017). *Supervising child protection practice: What works? An evidence informed approach*, Springer International Publishing. doi: 10.1007/978-3-319-5003

McWey, L., Cui, M., & Pazdera, A. (2010). Changes in Externalizing and Internalizing Problems of Adolescents in Foster Care. *Journal of Marriage and Family, 72*, 1128-40. doi: 10.1111/j.1741-3737.2010.00754.x.

Mihalic, S. (2004). The importance of implementation fidelity, *Emotional & Behavioral Disorders in Youth,* 4, 83-105.

Minuchin, S. (1974). *Families and family therapy*. Cambridge, MA: Harvard University Press.

Montgomery, P., Underhill, K., Gardner, F., Operario, D. & Mayo-Wilson, E. (2013). The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of Clinical Epidemiology*, *66*, 874-82. doi:10.1016/j.jclinepi.2013.03.006

Montgomery, P., Grant, S., Hopewell, S., Macdonald, G., Moher, D., Michie, S., & Mayo-Wilson, E. (2013). Protocol for CONSORT-SPI: An Extension for Social and Psychological Interventions. *Implementation Science, 8*, 99. doi:10.1186/1748- 5908-8-99

Measuring Recidivism. (n.d.). Retrieved March 02, 2018, from https://www.nij.gov/topics/corrections/recidivism/Pages/measuring.aspx

Moore, D., Tetlock, P. E., Tanlu, L., & Bazerman, M. H. (2006). Conflicts of interest and the case of auditor independence: moral seduction and strategic issue cycling. Academy of Management Review, 31, 10– 29. doi: 10.2307/20159182

Morgan, P. L., Staff, J., Hillemeier, M. M., Farkas, G., & Maczuga, S. (2013). Racial and ethnic disparities in ADHD diagnosis from kindergarten to eighth grade. *Pediatrics, 132*(1), 85–93. doi:10.1542/peds.2012-2390

Mulvey, E. P., Arthur, M. W., & Reppucci, N. D. (1993). The prevention and treatment of juvenile delinquency: A review of the research. *Clinical Psychology Review, 13*, 133-167. doi:org/10.1016/0272-7358(93)90038-N

Muncie, J. (2015). *Youth and Crime*. London, UK: Sage.

Newman, M., Vigurs, C., Perry, A. E., Hallam, G., Schertler, E. P. V., Johnson, M., & Wall, R. (2012). *A systematic review of selected interventions to reduce juvenile re-offending.* Research Evidence in Education Library. London: EPPI-Centre, Social Science Research

Unit, Institute of Education, University of London. Retrieved from http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3360

Oxman, A. D., & Guyatt, G. (1991). Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology*, 44, 1271-78.

Oxman, A. D. & Guyatt, G. H. (1993) The science of reviewing research. *Annals of the New York Academy of Sciences, 703*, 125-34. doi.org/10.1111/j.1749-6632.1993.tb26342.x

Pappadopulos, E., Woolston, S., Chait, A., Perkins, M. B., Connor, D. F., & Jensen, P. S. (2006). Pharmacotherapy of aggression in children and adolescents: efficacy and effect size. *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Academie canadienne de psychiatrie de l'enfant et de l'adolescent*, *15*(1), 27–39.

Petrosino, A., Turpin-Perosino, C., Hollis-Peel, M. E., Lavenberg, J. G. (2013) "Scared Straight" and other juvenile awareness programs for preventing juvenile delinquency. *Cochrane Database of Systematic Reviews*, CD002796. doi: 10.1002/14651858.CD002796.pub2.

Petterutti, A., Walsh, N, & Valesquez, T. (2009) *The costs of confinement: Why good juvenile justice policies make good fiscal sense*. Retrieved from http://www.justicepolicy.org/research/78

Petticrew, M. & Roberts, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology & Community Health, 57,* 527-29.

Pfeiffer, S. I., & Strzelecki, S. C. (1990). Inpatient psychiatric treatment of children and adolescents: A review of outcome studies. *Journal of the American Academy of Child & Adolescent Psychiatry, 29*(6), 847-853.http://dx.doi.org/10.1097/00004583-199011000-00001

Pieper, D., Beuchter, R., Jerinic, P., & Eikermann, M. (2012). Overviews of reviews often have limited rigor: A systematic review. *Journal of Clinical Epidemiology*, 65, 1267-1273. doi: 10. 1016/j.clinepi.2012.06.015

Pieper, D., Antoine, S. L., Neugebauer, E. A., & Eikermann, M. (2014). Up-to-dateness is often neglected in overviews: A systematic review. *Journal of Clinical Epidemiology, 67*, 1302-1308. doi: 10.1016/j.jclinepi.2014.08.008

Pottick, K., Hansell, S., Gaboda, D., & Gutterman, E. (1993). Child and adolescent outcomes of inpatient psychiatric services: A research agenda. *Children and Youth Services Review. 15*. 371-384. 10.1016/0190-7409(93)90010-7.

Prinz, R. J. (1998). Conduct Disorders. In S. B. Alan & H. Michel (Eds.), *Comprehensive Clinical Psychology* (pp. 527-538). Oxford: Pergamon Press.

Prochnow, J.E., DeFronzo, J.V. (1997). The Impact of Economic and Parental Characteristics on Juvenile Misconduct. *Journal of Emotional and Behavioral Disorders, 5*(2), 119-124. https://doi.org/10.1177/106342669700500206

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P.,& Hensley, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health, 38,* 65-76. doi: 10.1007/s10488-010-0319-7

Rohde, P., Waldron, H. B., Turner, C. W., Brody, J. L., & Jorgensen, J. (2012). Sequenced versus integrated treatment for adolescents with comorbid depressive and substance use disorders. Unpublished manuscript.

Robbins, M. S., Alexander, J. F., & Turner, C. W. (2000). Disrupting defensive family interactions in family therapy with delinquent adolescents. *Journal of Family Psychology, 14*, 688–701. https://doi.org/10.1037/0893- 3200.14.4.688

Robbins, M. S., Alexander, J. F., Newell, R. M., & Turner, C. W. (1996). The immediate effect of reframing on client attitude in family therapy. *Journal of Family Psychology, 10*, 28-34. http://dx.doi.org/10.1037/0893-3200.10.1.28

Robbins, M. S., Alexander, J. F., Turner, C. W. & Hollimon, A. (2016). Evolution of Functional Family Therapy as an Evidence-Based Practice for Adolescents with Disruptive Behavior Problems. *Family Process, 55*, 543–557. doi:10.1111/famp.12230

Robbins, M. S. & Turner, C. W. (2018). Call for revision of Weisman and Montgomery's review of Functional Family Therapy. *Research on Social Work, 29*, 347-57. doi.org/10.1177/1049731518814201

Rutter, M., Giller, H., & Hagell, A. (1998). *Antisocial behavior by young people*. New York, NY: Cambridge University Press.

Ryan, T., Leversee, T., & Lane, S. (2011). *Juvenile sexual offending: Causes, consequences and correction*. Hoboken, NJ; John Wiley & Sons.

Sackett D. L., Rosenberg W. M. C., Gray J. A. M., Haynes R. B., & Richardson W. S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ, 312*:71-72. 312. doi: https://doi.org/10.1136/bmj.312.7023.71

Sackett, D. L, Wennberg, J. E. (1997). Choosing the best research design for each question. It's time to stop squabbling over the "best" methods. *BMJ, 315*:1636.

Sacks, H. S., Berrier, J., Reitman, D., Ancona-Berk, V. A. & Chalmers, T. C. (1987) Meta-analysis of randomized controlled trials. *New England Journal of Medicine, 316*, 450-5. doi: 10.1056/NEJM198702193160806

Sanders, M. R., Kirby, J. N., Tellegen, C. L., & Day, J. J. (2014). The Triple P-Positive Parenting Program: A systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychology Review, 34*, 337-57. doi.org/10.1016/j.cpr.2014.04.003

Schulz, K. F., Altman, D. G., Moher, D. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ, 2010*;340:c332

Sektan, M., McClelland, M. M., Acock, A., et al. (2010). Relations between early family risk, children's behavioural regulation, and academic achievement. *Early Child Research Quarterly*, *25*, 464-479.

Sexton, T. L., & Alexander, J. F. (2000). Functional Family Therapy. Family Strengthening Series. *Juvenile Justice Bulletin* (pp. 1-9).

Sexton, T. L. (2002). Family-based empirically supported interventions. *The Counselling Psychologist, 30*, 238-261. doi:pdf/10.1177/0011000002302003

Sexton, T. L. (2004). Implementing Functional Family Therapy: Practical, theoretical, and empirical findings. *17th Annual RTC Conference*. Tampa, FL.

Sexton, T. L. (2004). *Levels of evidence for the models and mechanisms of therapeutic change in family and couples therapy*. NY, New York: Wiley.

Sexton, T. L., Ridley, C. R. and Kleiner, A. J. (2004), Beyond common factors: Multilevel process models of therapeutic change in marriage and family therapy. *Journal of Marital and Family Therapy, 30*: 131-149. doi:10.1111/j.1752-0606.2004.tb01229.x

Sexton, T. L. (2009). Functional family therapy: Traditional theory to evidence-based practice. In J. H. Bray & M. Stanton (Eds.), *The Wiley-Blackwell handbook of family psychology* (pp. 327-340). Hoboken, NJ: Wiley-Blackwell.

Sexton, T. & Turner, C. W. (2010). The effectiveness of Functional Family Therapy for youth with behavioral problems in a community practice setting. *Journal of Family Psychology, 24*, 339-48. doi: 10.1037/a0019406.

Shadish, W. R., Montgomery, L. M., Wilson, P., Wilson, M. R., Bright, I., & Okwumabua, T. (1993). Effects of family and marital psychotherapies: A meta-analysis. *Journal of Consulting and Clinical Psychology, 61*(6), 992-1002.

Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, Mass: Houghton Mifflin.

Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., & Bouter, L. M. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, *7*, 10. doi:10.1186/1471-2288-7-10

Sholevar, E., Baron, D. A., Aussetts, M. A. & Spiga, R. (2010). Treatment adherence in family therapy with youth as a predictor of time to arrest. *The American Journal of Family Therapy, 38*, 1-15. doi: 10.1080/01926180902754695

Sigrid, J., Qais, A. & Zepeda, V. (2013). Effectiveness and implementation of evidence-based practices in residential care settings. *Child and Youth Services Review, 35*, 642-656. doi:10.1016/j.childyouth.2013.01.007

Slesnick, N., & Prestopnik, J. L. (2009) Comparison of family therapy outcome with alcohol-abusing, runaway adolescents. Journal of Marital and Family Therapy, 35, 255-277. doi: 10.1111/j.1752-0606.2009.00121.x

Snyder, H. N., Sickmund, M. *Juvenile offenders and victims: 2006 National report*. 2006 Mar; Retrieved from http://www.ojjdp.gov/ojstatbb/nr2006/

Stanton, M. D. & Shadish, W. R. (1997). Outcome, attrition, and family–couples treatment for drug abuse: A meta-analysis and review of the controlled, comparative studies. *Psychological Bulletin, 122*, 170–191.

Tanner-Smith, E. E., Wilson, S. J., & Lipsey, M. W. (2013). The comparative effectiveness of outpatient treatment for adolescent substance abuse: A meta-analysis. *Journal of Substance Abuse Treatment, 44*, 145-158. doi:10.1016/j.jsat.2012.05.006

Tarolla, S. M., Wagner, E. F., Rabinowitz, J., & Tubman, J. G. (2002). Understanding and treating juvenile offenders: A review of current knowledge and future directions. *Aggression and Violent Behavior, 7*, 125-143. doi:org/10.1016/S1359-1789(00)00041-0

Thapar, A., Pine, D. S., Leckman, J. F., Scott, S., Snowling, M. J., & Taylor, E. (Eds.) (2015). *Rutter's child and adolescent psychiatry* (6th edition). West Sussex, UK: John Wiley & Sons Ltd.

Tolan, P. H. C. R. E. (1986). Family therapy with juvenile delinquents: A critical review of the literature. *Family Process, 25*, 619-650.

Trentacosta, C. J., Hyde, L. W., Goodlett, B. D., & Shaw, D. S. (2013). Longitudinal prediction of disruptive behavior disorders in adolescent males from multiple risk domains. *Child psychiatry and human development*, *44*, 561–72. doi:10.1007/s10578-012-0349-3

Ministry of Justice. (n.d.) Retrieved from https://www.gov.uk/government/organisations/ministry-of-justice

Office of Juvenile Justice and Delinquency Prevention (OJJDP). (n.d.). Retrieved from https://ojp.gov/about/offices/ojjdp.htm

Uttley, L. & Montgomery, P. (2017). The influence of the team in conducting a systematic review. *Systematic Reviews, 6*, 149. doi: 10.1186/s13643-017-0548-x.

van Urk, F. C. (2016). Function-focused implementation fidelity for complex interventions: The case of Studio Schools. Unpublished doctoral dissertation, University of Oxford, Oxford, UK.

Vaughn, M. G., & Howard, M. O. (2004). Adolescent Substance Abuse Treatment: A Synthesis of Controlled Evaluations. *Research on Social Work Practice, 14*, 325-335. doi:10.1177/1049731504265834

Waldron, H. B., Slesnick, N., Brody, J. L., Turner, C. W., & Peterson, T. R. (2001). Treatment outcomes for adolescent sub- stance abuse at 4- and 7-month assessments. *Journal of Consulting and Clinical Psychology, 69,* 802–813.

Waldron, H. B., & Turner, C. W. (2008). Evidence-based psychosocial treatments for adolescent substance abuse. *Journal of Clinical Child and Adolescent Psychology, 37*, 238–261. doi:10.1080/ 15374410701820133

Walker, D. F., McGovern, S. K., Poey, E. L., & Otis, K. E. (2004). Treatment effectiveness for male adolescent sexual offenders: A meta-analysis and review. *Journal of Child Sexual Abuse, 13,* 281–293.

Webb, S. (2001). Some considerations on the validity of evidence-based social practice in social work. *British Journal of Social Work, 31*, 57-79.

Webster-Stratton, C. (2006). *The Incredible Years*. Toronto, ON; Umbrella Press

Weerasekera, P. (2013). The state of psychotherapy supervision: recommendations for future training. *International Review of Psychiatry, 25*, 255-264. doi: 10.3109/09540261.2013.769431

Weisman, C. B., & Montgomery, P. (2018). Functional Family Therapy (FFT) for behavior disordered youth Aged 10–18: An overview of reviews. *Research on Social Work Practice, 29*, 333-46. doi.org/10.1177/1049731518792309

Weisman, C. B. & Montgomery, P. (2019). Response to: A critical review and call for revision of Weisman and Montgomery's review of Functional Family Therapy by Michael S. Robbins, Charles Turner. *Research on Social Work Practice, 29*, 358-360. doi.org/10.1177/1049731518824412

Weisman, C. B. & Montgomery, P. (2019). Implementation issues in Functional Family Therapy: A narrative analysis of the evidence. *Research on Social Work Practice* sagepub.com/journals-permissions. doi: 10.1177/1049731519867432

Weisman, C. B. (2019). Functional Family Therapy (FFT) for behaviour disordered youth in the greater New Orleans area post-Katrina: a reflection on implementation elements and concerns. Unpublished dissertation, University of Birmingham, Birmingham, UK.

Weisz, J. R., Hawley, K. M., & Jensen Doss, A. (2004). Empirically tested psychotherapies for youth internalizing and externalizing problems and disorders. *Child and Adolescent Psychiatric Clinics of North America, 13*, 729-815. doi:10.1016/j.chc.2004.05.006

White, S., Frick, P., Lawing, K., & Bauer, D. (2013). Callous-Unemotional Traits and Response to Functional Family Therapy in Adolescent Offenders. Behavioral Sciences & the Law. 31. 10.1002/bsl.2041.

Wilson, S. J., & Lipsey, M. W. (2000). Wilderness challenge programs for delinquent youth: A meta-analysis of outcome evaluations. *Evaluation and Program Planning, 23*(1), 1-12.http://dx.doi.org/10.1016/S0149-7189(99)00040-3

Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine, 33*, S130-S143. https://doi.org/10.1016/j.amepre.2007.04.011

Wintersteen, M. B., Mensinger, J. L., & Diamond, G. S. (2005). Do gender and racial differences between patient and therapist affect therapeutic alliance and treatment retention in adolescents? *Professional Psychology: Research and Practice, 36*, 400-08.

Woolfenden, S. R., Williams, K., & Peat, J. K. (2002). Family and parenting interventions for conduct disorder and delinquency: a meta-analysis of randomised controlled trials. *Archives of Disease in Childhood*, *86*(4), 251–256. doi:10.1136/adc.86.4.251

Woolfenden, S., Williams, K., & Peat, J. (2002). Family and parenting interventions for conduct disorder and delinquency: a meta-analysis of randomised controlled trials. *Archives of Disease in Childhood, 86*, 251-56. doi: 10.1136/adc.86.4.251

World Health Organization. ICD-10. (n.d.). Retrieved from https://www.who.int/